12-13-2012 12:00 AM

# Functional Genomic Investigation of Aromatic Aminotransferases Involved in Ephedrine Alkaloid Biosynthesis in Ephedra Sinica (Stapf)

Korey G. Kilpatrick
*The University of Western Ontario*

Supervisor
Dr. Frédéric Marsolais
*The University of Western Ontario*

Graduate Program in Biology
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

# FUNCTIONAL GENOMIC INVESTIGATION OF AROMATIC AMINOTRANSFERASES INVOLVED IN EPHEDRINE ALKALOID BIOSYNTHESIS IN *EPHEDRA SINICA* (STAPF)

(Spine title: Aminotransferases Involved in Ephedrine Alkaloid Biosynthesis)

(Thesis format: Monograph)

by

Korey Kilpatrick

Graduate Program in Biology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

# CERTIFICATE OF EXAMINATION

Supervisor

_____
Dr. Frédéric Marsolais

Co-Supervisor

_____
Dr. Norman Hüner

Supervisory Committee

_____
Dr. Mark Bernards

_____
Dr. Abdelali Hannoufa

Examiners

_____
Dr. Abdelali Hannoufa

_____
Dr. Susanne Kohalmi

_____
Dr. Denis Maxwell

The thesis by

**Korey Kilpatrick**

entitled:

Functional genomic investigation of aromatic aminotransferases involved in ephedrine alkaloid biosynthesis in *Ephedra sinica* (Stapf).

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science

_____ _____
Date Chair of the Thesis Examination Board

# Abstract

*Ephedra sinica* (Ephedraceae) is a broom-like shrub cultivated in arid regions of China, Korea and Japan. This medicinal plant accumulates large amounts of the ephedrine alkaloids, including (*S*)-cathinone, (1*R*,2*S*)-norephedrine, (1*R*,2*S*)-ephedrine and (1*S*,2*S*)-pseudoephedrine in its aerial tissues. These analogues of amphetamine mimic adrenaline and stimulate the sympathetic nervous system. While much is known about their pharmacological properties, the biological mechanisms by which they are synthesized remains largely unknown. A functional genomics platform was established in order to investigate alkaloid biosynthesis. RNA was extracted from *Ephedra sinica* stems and sequenced by Illumina HiSeq2000 next-generation sequencing. Candidate biosynthetic enzymes were obtained from this EST collection based on similarity to characterized enzymes with similar functions. This portion of the collaborative study is focused on identifying aminotransferase enzymes involved in alkaloid biosynthesis. Two functional aromatic aminotransferase enzymes, possibly involved in ephedrine biosynthesis have been identified. One exhibiting much greater enzymatic activity was chosen for characterization.

**Keywords:** Aminotransferase, enzyme kinetics, *Ephedra sinica,* ephedrine alkaloids, functional genomics, medicinal plants, next-generation sequencing, secondary metabolism

# Acknowledgments

I would like to sincerely thank my supervisor, Dr. Frédéric Marsolais for his support and guidance over the past few years; especially for providing an intellectually inspiring environment to further my academic studies and always being available for research related discussion. I would like to thank my co-supervisor, Dr. Norm Hüner for his guidance and helpful suggestions in preparing and editing my thesis. I would like to thank my advisory committee: Dr. Mark Bernards and Dr. Ali Hannoufa for their guidance, input on the direction of research and editing of my thesis. I also would like to thank Dr. Susanne Kohalmi for her assistance with ADT/PDT phylogenetic analysis.

I am very thankful to the research technicians and staff at Agriculture and Agri-Food Canada (London, Ontario) for their patience, insight and technical support: Aga Pajak, Sudhakar Pandurangan, Lisa Amyot, Ted Blazejowski and Alex Molnar. I am grateful for my fellow lab mates, friends and colleagues who were always available to assist with experiments, discuss one another's projects or to just go for coffer or lunch break: Nikita Eskin, Mehran Dastmalchi, Lyne Sabourin, Pat Yoon, Eridan Pereira, Danny Kim and Matt Clarke.

I am thankful for the Agriculture and Agri-Food Canada/Natural Science and Engineering Research Council/United States – Israel Binational Agricultural Research and Development Fund (BARD), Canada – Israel Cooperation in Agricultural Research Program, Phase II for funding my research.

Lastly, I would like to thank my family and friends, whose support, love and guidance have made me who I am today.

# Table of Contents

# List of Tables

# List of Figures

# List of Appendices

# List of Abbreviations

| | |
|---|---|
| ADT | Arogenate dehydratase |
| BLAST | Basic local alignment search tool |
| bp | Base pair(s) |
| BSA | Bovine serum albumin |
| cDNA | Complimentary DNA |
| CDSs | Coding DNA sequences |
| DEPC | Diethylpyrocarbonate |
| DNA | Deoxyribonucleic acid |
| DTT | Dithiothreitol |
| EDTA | Ethylenediaminetetraacetic acid |
| EST | Expressed sequence tag |
| GC-MS | Gas chromatography-mass spectrometry |
| GST | Glutathione S-transferase |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| Kbp | Kilobase pair(s) |
| $k_{cat}$ | Apparent unimolecular rate constant |
| $k_{cat}/K_m$ | Catalytic efficiency |
| $K_m$ | Michaelis-Menten constant |
| LB | Luria Broth |

| | |
|---|---|
| LIC | Ligation independent cloning |
| LSD | Least significant difference |
| MAGPIE | Automated genomics project investigation environment |
| $NAD^+$ | β-Nicotinamide adenine dinucleotide |
| NADH | β-Nicotinamide adenine dinucleotide, reduced |
| $OD_{600}$ | Optical density at 600 nm |
| ORF | Open reading frame |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| PDT | Prephenate dehydratase |
| PLP | Pyridoxal-5'-phosphate |
| PVP-40 | Polyvinylpyrrolidone |
| PVPP | Polyvinylpolypyrrolidone |
| RNA | Ribonucleic acid |
| rRNA | Ribosomal RNA |
| SDS-PAGE | Sodium dodecyl sulfate – polyacrylamide gel electrophoresis |
| tRNA | Transfer RNA |
| UV | Ultra-violet |
| $V_{max}$ | Maximum reaction rate |
| $\varepsilon_{331}$ | Extinction coefficient (at 331 nm) |

# Chapter 1. Introduction

## 1.1 *Ephedra sinica*

*Ephedra sinica* Stapf (Ephedraceae) is a perennial, broom-like shrub cultivated in arid regions of China, Korea and Japan (Bensky *et al.*, 1986). It belongs to the gnetophyta division of gymnospermic plants and is related to the conifers (Rydin and Korall, 2009). This traditional Chinese medicinal plant has been utilized by humans for over 5,000 years; *Ma Huang*, a tea containing dried *Ephedra sinica* stems was utilized for the treatment of asthma, fever, coughs and lack of sweating, and to promote urination and reduce edema (Bensky *et al*., 1986; Bruneton, 1995). All aerial tissues of this plant produce the ephedrine alkaloids, with the stems accumulating the highest amounts, up to 3% of dry weight (Krizevski *et al*., 2010). This species was chosen because it produces all of the ephedrine alkaloids and in large amounts compared to other *Ephedra* species (Cui *et al*., 1991).

## 1.2 Ephedrine Alkaloids and Biosynthetic Pathway

### *1.2.1 Ephedrine Alkaloids*

Alkaloids are a large and diverse group of compounds which contain a nitrogen atom in a heterocyclic backbone (Facchini and De Luca, 2008). The ephedrine alkaloids are a class of phenylpropylamino alkaloids which are analogues of amphetamine and as such they mimic the actions of adrenaline in mammalian hosts by stimulating the sympathetic nervous system (Bruneton, 1995). These pharmacoactive stimulants, including (*1S,2S*)-norpseudoephedrine, (*1R,2S*)-norephedrine, (*1S,2S*)-pseudoephedrine, (*1R,2S*)-ephedrine and (*S*)-cathinone are collectively known as the ephedrine alkaloids. In

modern medicine (*1R,2S*)-norephedrine and (*1R,2S*)-ephedrine are used to treat acute asthma attacks, rhinitis and sinusitis because of their vasoconstrictive properties (Bruneton, 1995). (*1R,2S*)-Ephedrine is also utilized to counteract anesthesia and overdoses of depressant drugs as well as treatment for bronchial asthma and various allergies (Lewis and Elvin-Lewis, 1977). (*1S,2S*)-Pseudoephedrine is a widely used nasal decongestant (Bruneton, 1995). Controversially, (*1R,2S*)-ephedrine is used as a weight-loss or exercise supplement and a number of deaths have been attributed to its use (Haller and Benowitz, 2000). The true ephedrine alkaloids, containing nitrogenous rings, (*2S,4S,5S*)-2,4-dimethyl-5-phenyloxazolidine, (*2S,4S,5R*)-2,4-dimethyl-5-phenyloxazolidine, (*2S,4S,-5S*)-2,3,4-trimethyl-5-phenyloxazolidine and (*2S,4S,5R*)-2,3,4-trimethyl-5-phenyl-oxazolidine are formed from the spontaneous cyclization of (*1S,2S*)-norpseudoephedrine, (*1R,2S*)-norephedrine, (*1S,2S*)-pseudoephedrine, (*1R,2S*)-ephedrine, respectively (Krizevski *et al*., 2010).

Pseudoephedrine and ephedrine are extensively used pharmaceuticals globally. In their 2008 Assessment of Annual Needs, the American Drug Enforcement Agency estimated the annual global requirement for pseudoephedrine and ephedrine at 511 and 140 tons, respectively (http://www.deadiversion.usdoj.gov/fed_regs/quotas/2008/-fr0623.htm) (Leonhart, 2008).

(*S*)-Cathinone is the main psychostimulant found in *Catha edulis* (khat) leaves, chewed at social gatherings in the Middle East and East Africa. It is a putative precursor of the ephedrine alkaloids in *Ephedra sinica* (Krizevski *et al*., 2010). Khat is a taxonomically unrelated plant, yet it appears to produce many of the ephedrine alkaloids via the same intermediates (Krizevski *et al*., 2008). This suggests the ephedrine alkaloid

pathway is a result of convergent evolution. To date, four plant genus' which produce the ephedrine alkaloids have been identified: *Ephedra*, *Catha* (Hagel *et al*., 2011; Krizevski *et al.*, 2007; Lewis and Elvin-Lewis, 1977), *Pinellia* (Bensky *et al*., 1986; Oshio *et al.*, 1978) and *Sida* (Khatoon *et al*., 2005). While much is known about the pharmacological properties of these stimulants, the biological mechanisms by which they are synthesized remains largely unknown.

### *1.2.2 Ephedrine Alkaloid Biosynthesis*

A proposed biosynthetic pathway for the production of the ephedrine alkaloids in *Ephedra* species is presented in Figure 1.1 (Hagel *et al*., 2012). Early radiolabelling experiments using khat leaves determined that L-phenylalanine is the initial precursor of the ephedrine alkaloids (Leete, 1958). Those findings were later confirmed using *Ephedra distachya* stems (Yamasaki *et al*., 1969; Yamasaki *et al*., 1973). However, only the $C_6$-$C_1$ subunits of phenylalanine, the benzylic ring, are incorporated into ephedrine (Yamasaki *et al*., 1969).

L-Phenylalanine is first converted to cinnamic acid by phenylalanine ammonia lyase (Leete, 1958; Okada *et al*., 2008; Yamasaki *et al*., 1969) and then further catalyzed to benzaldehyde or benzoic acid (Yamasaki *et al*., 1973). Three possible methods of chain shortening have been proposed: non-β-oxidative and CoA independent, β-oxidative and CoA dependent and a non-β-oxidative, CoA dependent method believed to occur in some plants (Boatright *et al*., 2004).

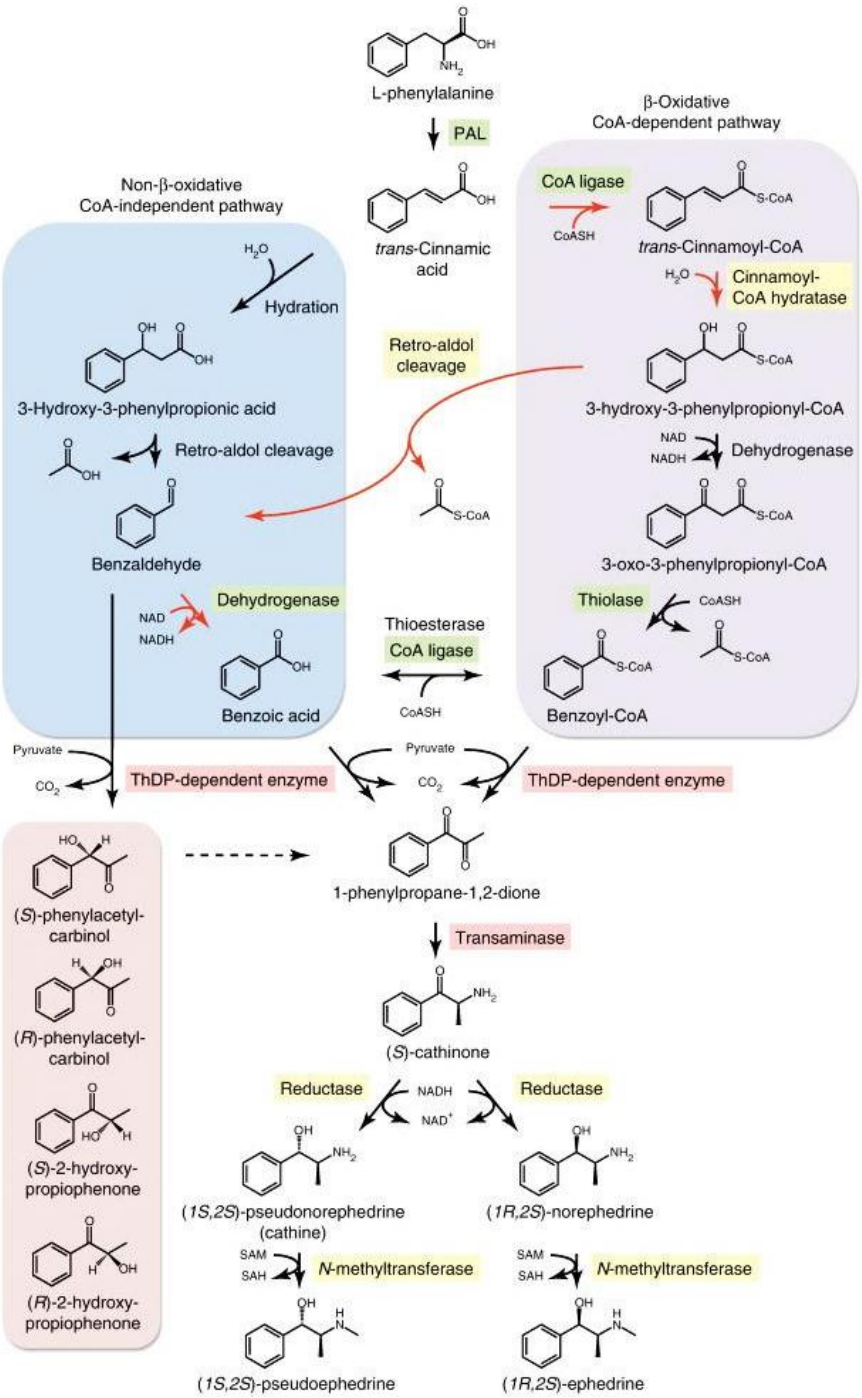


**Figure 1.1 Proposed biosynthetic pathway leading from L-phenylalanine to the ephedrine alkaloids.** A CoA-independent, non-β-oxidative pathway of side-chain shortening is shown in blue, a CoA-dependent, β-oxidative route is shown in purple. Red arrows indicate an alternative CoA-dependent, non-β-oxidative route. Benzoic acid or benzoyl-CoA undergoes condensation with pyruvate, putatively catalyzed by a ThDP-dependent enzyme. Benzaldehyde may also undergo carboligation, producing numerous possible products (pink), a reaction observed in numerous microbes. 1-Phenylpropane-1,2-dione is thought to be an immediate precursor to (*S*)-cathinone via transamination. (*S*)-Cathinone is reduced to (*1S,2S*)-norpseudoephedrine and (*1R,2S*)-norephedrine. N-Methylation then occurs producing (*1S,2S*)-pseudoephedrine and (*1R,2S*)-ephedrine. Activity has been detected for enzymes highlighted in yellow, and corresponding cDNAs are available for enzymes highlighted in green. Enzymes highlighted in pink have not been previously detected. Abbreviations: CoA, Coenzyme A; PAL, phenylalanine ammonia lyase; ThDP, thiamine diphosphate (Reproduced from Hagel *et al*., 2012 with permission from Elsevier).

Using *Ephedra gerardiana sikkimensis* it was later shown that pyruvic acid provides the $C_2$ unit of ephedrine and that it is incorporated with benzoic acid (or its CoA thioester), not benzaldehyde (Grue-Sørensen and Spenser, 1989, 1993, 1994). However, benzaldehyde could easily be oxidized, non-enzymatically, prior to incorporation (Nierop Groot and De Bont, 1998). In contrast, it was recently shown in cell-free extracts of *Ephedra sinica* and *Ephedra foeminea* stems that benzaldehyde and pyruvate are ligated (termed benzaldehyde carboxyligase activity) to form (*R*)-phenylacetylcarbinol, (*S*)-phenylacetylcarbinol, (*R*)-2-hydroxy-propiophenone (*S*)-2-hydroxypropiophenone and 1-phenylpropane-1,2-dione (Krizevski *et al*., 2012). Although (*R*)- and (*S*)-phenylacetylcarbinol were not incorporated into the ephedrine alkaloids in *Ephedra gerardiana sikkimensis* (Grue-Sørensen and Spenser, 1994) it is possible for these intermediates to be transaminated directly into (*1R,2S*)-norephedrine and (*1S,2S*)-pseudoephedrine, respectively. This is similar to the industrial production of ephedrine where benzaldehyde is fed to yeast and undergoes pyruvate decarboxylase-catalyzed carboligation to form (*R*)-phenylacetylcarbinol which is extracted and subjected to reductive amination, producing (*1R,2S*)-ephedrine (Hagel *et al*., 2012).

1-Phenylpropane-1,2-dione is thought to be transaminated to form (*S*)-cathinone, both of which were incorporated into the ephedrine alkaloids in *Ephedra gerardiana sikkimensis* (Grue-Sørensen and Spenser, 1994). The amino donor as well as the biochemical conditions of this reaction is currently unknown. This reaction may occur in a similar manner to that of tyrosine or aromatic aminotransferase which utilize L-glutamate as amino donor and keto acid derivatives of L-tyrosine (4-hydroxyphenylpyruvate) and L-phenylalanine (phenylpyruvate) as amino acceptors in the reverse direction (Gonda *et al.*, 2010; Lee and Facchini, 2011; Prabhu and Hudson,

2010; Riewe *et al*., 2012). 1-Phenylpropane-1,2-dione is the first $C_6$-$C_3$ intermediate and as such is considered to be the first committed precursor in the pathway. The reduction of (*S*)-cathinone to yield a nearly 1:1 ratio of (*1S,2S*)-norpseudoephedrine and (*1R,2S*)-norephedrine was shown to be remarkably efficient (Grue-Sørensen and Spenser, 1994). Coupled with the fact that (*R*)- and (*S*)- phenylacetylcarbinol were not incorporated into the ephedrine alkaloids, this suggests the route to the ephedrine alkaloids is through (*S*)-cathinone and not one of the phenylacetylcarbinol isomers. However, phenylacetylcarbinol should still be considered as a possible substrate in an aminotransferase reaction during ephedrine alkaloid biosynthesis due to the recent report of benzaldehyde carboxyligase activity (Krizevski *et al.*, 2012).

Methylation of the norephedrines' nitrogenous amine group to produce (*1S,2S*)-pseudoephedrine and (*1R,2S*)-ephedrine was revealed over 50 years ago (Shibata *et al*., 1957) and recently confirmed in *Ephedra sinica* cell-free extracts (Krizevski *et al.*, 2012). Although many of the intermediates in the ephedrine alkaloid biosynthetic pathway have been identified, the specific enzymes and biochemical conditions which facilitate the production of these precursors and alkaloids themselves are largely unknown. The only step to have been characterized at the molecular level in *Ephedra sinica* is the initial deamination of L-phenylalanine by phenylalanine ammonia lyase (Okada *et al*., 2008).

### *1.2.3 Timeline of Ephedrine Alkaloid Production*

With the use of chiral GC-MS, 1-phenylpropane-1,2-dione and (*S*)-cathinone were recently detected in *Ephedra sinica* stem tissue for the first time. However, both precursors were only detected in young *Ephedra sinica* stems and not detected in mature tissue. Mature stems accumulate up to seven-fold greater amounts of the ephedrine

alkaloids compared to young stems. There is also large variation in alkaloid composition and content between different accessions of *Ephedra sinica* plants grown under identical conditions (Krizevski *et al*., 2010). Taken together this provides a developmental timeline for the production of the ephedrine alkaloids and strongly suggests alkaloid profiles are determined by genetic factors. Therefore, functional genomics will be utilized to identify candidate biosynthetic enzymes involved in ephedrine alkaloid biosynthesis. This study is focused on elucidating possible aminotransferase-like enzymes responsible for the production of L-phenylalanine and (*S*)-cathinone.

## 1.3 Aminotransferase Enzymes

### *1.3.1 Aminotransferase Reaction Mechanism*

An aminotransferase (also known as a transaminase) is an enzyme that catalyzes the reversible transfer of an amino group from an amino donor, typically an amino acid, to the carbonyl group of an amino acceptor, typically a keto acid. These enzymes are involved in a number of essential biological processes including anabolism and catabolism of amino acids, nitrogen assimilation, vitamin biosynthesis, plant stress responses and secondary metabolism (Givan, 1980; Liepman and Olsen, 2004).

Aminotransferases are part of the pyridoxal-5'-phosphate (PLP) dependent class of enzymes since they require PLP as a cofactor. PLP is covalently bound through an imine linkage to the ε-amino group of the conserved lysine residue in the active site of the enzyme (Mehta *et al*., 1993). Most aminotransferase enzymes are able to accept multiple substrates and employ a ping-pong reaction mechanism (Hirotsu *et al*., 2005). First, the amino donor enters the active site and becomes bound to PLP at its alpha carbon forming

a Schiff base linkage known as the aldimine. The lysine residue in the active site de-protonates the alpha carbon creating a quinonoid intermediate. The $C_4$ of PLP is then re-protonated by the lysine residue generating a ketimine intermediate which is hydrolyzed to release the oxo-acid derivative of the amino donor. This leaves the enzyme in the pyridoxamine phosphate form, now ready to donate the bound amino group to a carbonyl group on the amino acceptor which would regenerate the PLP form of the enzyme (Eliot and Kirsch, 2004; Hirotsu *et al*., 2005; Kirsch *et al*., 1984; Kochhar and Christen, 1992; Kochhar *et al*., 1987; Kuramitsu *et al*., 1987).

*1.3.2 Aromatic Aminotransferases*

Plant aromatic aminotransferases are able to catalyze reactions using L-tyrosine, L-phenylalanine and L-tryptophan as amino donors in the forward direction. In the reverse direction, L-glutamate is typically the preferred amino donor. They are either named tyrosine aminotransferase or aromatic aminotransferase depending on their preferred substrate(s) but are frequently able to use all three aromatic amino acids as substrates. They typically utilize α-ketoglutarate as the preferred amino acceptor in the forward direction but have also been shown to accept oxaloacetate and pyruvate as substrates. In the reverse direction, keto acid derivatives of the aforementioned aromatic amino acids are amino acceptors (Hirata *et al*., 2012; Prabhu and Hudson, 2010; Ziegler *et al*., 2009). Aminotransferases are said to be promiscuous enzymes because of their ability to accept multiple amino donors and acceptors as substrates.

Recently, a number of aminotransferases which utilize aromatic substrates have been identified in plants. These enzymes have roles in varying biological processes. Two tyrosine aminotransferases, one with a role in tocopherol biosynthesis, have recently been

identified and characterized from *Arabidopsis thaliana* (Prabhu and Hudson, 2010; Riewe *et al*., 2012). Another tyrosine aminotransferase with a role in benzylisoquinoline alkaloid biosynthesis was recently characterized in the opium poppy (Lee and Facchini, 2011). An aromatic aminotransferase involved in volatile aroma production in musk melon fruit was recently identified (Gonda *et al*., 2010) and an aminotransferase that prefers to catabolize phenylalanine and named aromatic or phenylalanine aminotransferase, was recently characterized from rose petal protoplasts (Hirata *et al.*, 2012). These plant aminotransferases which are able to utilize aromatic substrates will aid in the identification of candidate aminotransferase-like enzymes putatively involved in the biosynthesis of L-phenylalanine and (*S*)-cathinone.

### 1.4 Rationale, Hypothesis and Objectives

This study investigates the role of aminotransferase-like enzymes in the ephedrine alkaloid biosynthetic pathway of *Ephedra sinica*. As previously mentioned, an aminotransferase reaction is proposed to take place using 1-phenylpropane-1,2-dione as a substrate to synthesize (*S*)-cathinone. The initial precursor of the ephedrine alkaloids, L-phenylalanine could also be produced, in part from transamination of phenylpyruvate (Maeda and Dudareva, 2012). With the recent emergence and decreasing costs involved in advanced "-omics" methodologies, researchers have the unprecedented ability to elucidate the biochemistry of pathways which produce high-value plant metabolites (Facchini and De Luca, 2008; Ziegler and Facchini, 2008). In order to investigate ephedrine alkaloid biosynthesis, a snapshot of the *Ephedra sinica* stem transcriptome was generated by massively parallel, next-generation sequencing of RNA extracted from young *Ephedra sinica* stems. This functional genomics platform was triaged for candidate

biosynthetic genes by screening for sequence similarity to previously identified plant aminotransferases. Candidate cDNAs were cloned and expressed in a heterologous system and assayed *in vitro* to determine their function.

It is hypothesized that the enzyme catalyzing the conversion of 1-phenylpropane-1,2-dione to (*S*)-cathinone will have significant sequence similarity to previously identified aromatic aminotransferases from plant species. This is due to the fact that 1-phenylpropane-1,2-dione and (*S*)-cathinone contain benzene rings and are structurally similar to the aromatic amino acids. It is also hypothesized that an aromatic aminotransferase is responsible for producing some of the L-phenylalanine utilized by the plant to produce the ephedrine alkaloids.

The objectives of this study are to (1) isolate RNA from *Ephedra sinica* which will be utilized for next-generation sequencing to assist in the development a functional genomics platform in order to investigate biosynthesis of the ephedrine alkaloids. (2) Characterize candidate cDNAs encoding aminotransferase-like enzymes. (3) Develop *in vitro* assays to detect recombinant enzyme activity. (4) Identify the enzyme catalyzing the conversion of 1-phenylpropane-1,2-dione to (*S*)-cathinone and (5) identify the enzyme catalyzing the conversion of phenylpyruvate to L-phenylalanine.

# Chapter 2. Materials and Methods

## 2.1 *Ephedra sinica* Plant Growth

*Ephedra sinica* seeds were purchased from Horizon Herbs (Williams, Oregon, USA) and Richters Herbs (Goodwood, Ontario, Canada). Seeds were sown in the spring and fall months in 50% sand, 50% Cactus Soil mixture (Hortibec) and grown in a greenhouse with at least 14 hours of light per day. Supplemental 220 V, high pressure sodium lighting was used. Seedlings were grown in individual trays until they reached approximately four inches in height when they were transferred to larger pots. Young seedlings were watered every-other day before transplanting, every four days (as needed to prevent damping off) after transplanting and fertilized with 20:20:20 (nitrogen-phosphorus-potassium) fertilizer once a week. 'Young' *Ephedra sinica* tissues used in this study were the freshly grown tips of the stems and light green in colour. 'Mature' tissues were the lower portions of stems and dark green in colour. After sample collection, fresh plant material was immediately frozen in liquid nitrogen and stored at -80°C until use. All chemicals and commercially available enzymes were purchased from Sigma-Aldrich unless otherwise stated.

## 2.2 RNA Extraction

Total RNA was extracted from *Ephedra sinica* stems following a modified protocol for extracting RNA from tissues containing high levels of procyanidins (Wang and Vodkin, 1994). Stems from multiple plants were ground to a fine powder with a mortar and pestle under liquid nitrogen. Approximately 500 mg of ground tissue was added to 2 mL of Basic RNA Extraction Buffer (100 mM Tris-HCl pH 9.0, 20 mM

sodium EDTA, 4% (w/v) sarkosyl, 200 mM sodium chloride, 16 mM dithiothreitol (DTT) and 16 mM mercaptobenzothiol) with 5% bovine serum albumin (BSA) and 10 mg/mL heparin and ground to a homogenate with a small amount (~2 g) of polyvinylpolypyrrolidone (PVPP), hydrated in Basic RNA Extraction Buffer. An additional 6.5 mL of Basic RNA Extraction Buffer was added and tissue was ground for one minute. Next, 100 µL of 20 mg/mL Proteinase K (BioShop Canada Inc.) was added and the crude extract incubated for 20 min at 37°C with gentle shaking (80 rpm). Solution was then centrifuged at 30,000 x *g* for 20 min at 4°C. The supernatant was transferred to a clean tube and extracted twice with an equal volume of saturated acidic phenol (pH 4.0). After each extraction, the phases were separated by centrifugation (30,000 x *g* for 20 min at 20°C). The top organic layer was then extracted with an equal volume of Sevag solution (24 parts chloroform to 1 part isoamyl alcohol) and the phases separated by centrifugation (as above). After the top organic layer was transferred to a clean tube, one-third of the volume of 8 M lithium chloride was added drop-wise while vortexing. The solution was placed at 4°C overnight to precipitate the RNA. Solutions were centrifuged at 25,000 x *g* for 20 min at 4°C and the supernatant discarded. The pellet was washed twice with 500 µL of 2 M lithium chloride, centrifuging at 25,000 x *g* for 5 min at 4°C after each wash. After resuspension in 100 µL of diethylpyrocarbonate (DEPC) treated water, 10 µL of 3 M sodium acetate (pH 5.2) and 300 µL of 95% (v/v) ethanol was added and the solution was placed at -20°C overnight. The pellet was collected by centrifugation (25,000 x *g* for 20 min at 4°C) and washed twice with 70% (v/v) ethanol. The RNA pellet was dried, resuspended in DEPC water and stored at -80°C until use. All glassware and utensils utilized for RNA extraction were baked overnight at 180°C or treated with RNase Away (Molecular BioProducts) to prevent RNA degradation. Guidelines for RNA

extraction and handling were followed as outlined in Current Protocols in Molecular Biology (Kingston, 2001).

**2.3 Next-Generation Sequencing, Bioinformatics and Gene Triage**

A sample of total RNA (10 μL of 2.3 μg/μL RNA with a 260/280 ratio of 2.10 and a 260/230 ratio of 2.19) extracted from young *Ephedra sinica* stems was sent for next-generation sequencing on the Illumina HiSeq 2000 RNA-seq platform (single channel, 108 base pair, paired-end reads). The sequencing was performed as per the manufacturers' specifications by the Massively Parallel Sequencing Unit of the McGill University and Génome Québec Innovation Centre. Expressed sequence tag (EST) data from the library was assembled using the Velvet algorithms for *de novo* short read assembly (Zerbino and Birney, 2008) and deposited in the MAGPIE (Automated Genomics Project Investigation Environment) portal at the University of Calgary, in collaboration with Dr. Christoph Sensen and the Genome Canada PhytoMetaSyn project. This data was mined for candidate genes in the ephedrine alkaloid biosynthetic pathway based on sequence similarity to previously characterized enzymes with similar functions. For cathinone aminotransferase/synthase, BLAST searches were performed using, as queries, aromatic aminotransferases from melon (GenBank accession number: ADC45389.1) (Gonda *et al*., 2010) and rose (AB669189.1) (Hirata *et al*., 2012), tyrosine aminotransferases from *Arabidopsis* (AED94050.1) (Prabhu and Hudson, 2010) and opium poppy (ADC33123.1) (Lee and Facchini, 2011), prephenate aminotransferase from petunia (ADM67557.1) (Maeda *et al*., 2011), tryptophan aminotransferase from *Arabidopsis* (AEE35079.1) (Stepanova *et al*., 2008) and aspartate aminotransferases from

pine (CAF31327.1) (de la Torre *et al.*, 2006) and alfalfa (AAB46610.1) (Udvardi and Kahn, 1991).

**2.4 Isolation and Cloning of Candidate cDNAs**

*Ephedra sinica* RNA was first treated with DNaseI (Ambion) following manufacturers' guidelines to remove any residual DNA. The RNA was then reverse transcribed into cDNA using Quanta qScript cDNA Super mix following manufacturers' guidelines and stored at -20°C until use.

Primers for each candidate gene (Appendix A1) were designed based on the longest deduced open reading frame (ORF) and sequence alignment to query genes. The general PCR reaction used *Pfx50* DNA polymerase (Invitrogen) and consisted of denaturing at 94°C for 2 min, 35 cycles of 94°C for 15 s, 60.5°C for 30 s and 68°C for 90 s, and a final extension at 68°C for 5 min. PCR products for Contigs 13244 and 12701 were cloned into pCR4Blunt-TOPO (Invitrogen) vector to confirm the proper sequence and then subcloned into pQE30 expression vector by restriction enzyme-mediated cloning using *BamHI* and *KpnI* restriction endonucleases (New England Biolabs) and T4 DNA ligase (Invitrogen) following manufacturers' guidelines. *Escherichia coli* (*E. coli*) strain XL10-Gold was used for cloning and expression of genes in the pQE30 vector. Contigs 4103 and 12559 were subcloned into the pPROEX HTb expression vector and transformed into BL21 (DE3) *E. coli* (Invitrogen) for heterologous protein expression.

Several candidate cDNAs (Contigs 12559, 287, 12253, 19895, 22464 and 798, and Singlet 4072) were cloned into the pLATE51 expression vector using aLICator ligation independent cloning (LIC) system (Fermentas) and transformed into BL21 (DE3)

*E. coli* for expression. The PCR reactions used *AccuPrime Pfx* DNA Polymerase and consisted of denaturing at 95°C for 2 min, 35 cycles of 95°C for 15 s, 55°C for 30 s and 68°C for 1 min 45 s, and a final extension at 68°C for 5 min. PCR products were cloned into the pLATE51 vector following manufacturers' guidelines and sequences were confirmed by Sanger sequencing.

Seven difficult to express cDNAs (Contigs 19156, 901, 4535, 18610, 16680, 29549 and 11412) were fused to glutathione *S*-transferase (GST). Their PCR products were cloned into pCR4Blunt-TOPO vector to confirm the proper sequence and then subcloned into the pGEX-6P-2 expression vector (GE Healthcare Life Sciences) by restriction enzyme-mediated cloning using *BamHI* and *SalI* (New England Biolabs) and T4 DNA ligase (Invitrogen) following manufacturers' guidelines. These constructs were also transformed into BL21 (DE3) *E. coli* for expression.

**2.5 Heterologous Expression of Recombinant Proteins**

Fresh overnight culture containing recombinant plasmid was grown in LB media and used to inoculate 100 or 500 mL of NZY media. Cultures were incubated at 37°C until an $OD_{600}$ between 0.5 and 0.6 was reached. Expression of the candidate genes was induced by the addition of 0.5 or 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and cells were cultured overnight at room temperature (Contigs 13244, 12701 and 4103, and Singlet 4072) or at 16°C (Contigs 12559, 287, 12253, 19895, 22464, 798, 19156, 901, 4535, 18610, 16680, 29549 and 11412). Following expression, cells were pelleted by centrifugation at 5,500 x *g* for 5 min at 4°C and frozen at -20°C to enhance lysis.

## 2.6 Affinity Chromatography for Purification of Recombinant Proteins

### *2.6.1 Purification of Contigs 13244, 12701, 4103, 12559, 287, 12253, (Singlet) 4072, 19895, 22464 and 798*

Pelleted cells were resuspended in 5 – 10 mL of Native Protein Lysis Buffer (50 mM sodium phosphate monobasic, 300 mM sodium chloride and 10 mM imidazole at pH 8.0), 1 mg/mL lysozyme was added and cells were incubated on ice for 30 min. A French Press (American Instrument Co. Inc.) was used to exert 1200 psi pressure on the cells to lyse them and enhance protein recovery. Sonication was also performed to break down genomic DNA and reduce viscosity. The lysate was then centrifuged at 15,500 x *g* for 45 min at 4°C to pellet cell debris. To the supernatant, 500 µL Ni-NTA agarose beads (Qiagen) were added to bind the hexahistidine tagged recombinant proteins. The supernatant and nickel beads were incubated at 4°C for 2 hours on a slow rotor. The Ni-NTA protein solutions were centrifuged at 3,500 x *g* for 3 min at 4°C and the supernatant discarded. The pellets were then washed 10 times with Native Protein Wash Buffer. Recombinant proteins were eluted from the beads by the addition of 1.25 mL Native Protein Elution Buffer (50 mM sodium phosphate monobasic, 300 mM sodium chloride and 250 mM imidazole at pH 8.0). The solution was incubated for 10 min at 4°C on a slow rotor and elution repeated with an additional 1.25 mL of Native Protein Elution Buffer. Buffer exchange was performed using a PD-10 column (GE Healthcare Life Sciences) following the manufacturers' gravity protocol. Recombinant proteins were concentrated using an Amicon Ultra-4 filter and stored at -80°C in 100 mM HEPES-NaOH (pH 8.2) with 1 mM sodium EDTA and 20% (v/v) glycerol.

*2.6.2 Purification of Contigs 19156, 901, 4535, 18610, 16680, 29549 and 11412*

Pelleted cells were dissolved in 5 – 10 mL of ice-cold PBS (137 mM sodium chloride, 2.7 mM potassium chloride, 10.1 mM sodium phosphate dibasic and 1.8 mM potassium phosphate monobasic at pH 7.4), 1 mg/mL lysozyme was added and cells were incubated on ice for 30 min. Lysates were then sonicated on ice, in short bursts (10 s burst followed by 10 s pause, repeated six times) to promote cell lysis. Triton X-100 was added to a 1% (v/v) final concentration and the samples were incubated for 30 min at 4°C with end-over-end rotation. Samples were then centrifuged at 15,000 x *g* for 30 min at 4°C to pellet cell debris. The supernatant was transferred to a fresh container containing glutathione Sepharose 4B slurry (GE Healthcare Life Sciences). A 50 µL bed volume was used per 100 mL of culture. Samples were incubated for 30 min at room temperature with end-over-end rotation. The medium with absorbed fusion protein was centrifuged at 500 x *g* for 5 min at 4°C and washed three times with 1 mL of ice-cold PBS. Proteins of interest were cleaved from the GST fusion protein using PreScission Protease (GE Healthcare Life Sciences) following manufacturers' guidelines. Samples were then centrifuged at 500 x *g* for 5 min at 4°C to sediment the glutathione sepharose. The supernatant containing protein of interest was transferred to a PD-10 column to exchange buffers following the manufacturers' gravity protocol. Recombinant proteins were concentrated and stored as described in 2.6.1.

Protein concentration was determined by Bio-Rad protein assay solution using BSA as a standard and following manufacturers' guidelines. SDS-PAGE was performed using Coomassie Brilliant Blue R250 (or G250) to stain proteins to confirm the proper protein had been purified by comparison to the predicted molecular mass. Molecular

masses and isoelectric points of recombinant proteins were predicted using ExPASy Compute pI/MW tool.

### 2.7 Protein Extraction from *Ephedra sinica* Stem Tissue

Young *Ephedra sinica* stems (20 g) were ground with a mortar and pestle under liquid nitrogen and with 1/10 the amount (2.0 g) of washed sea sand (Fisher Scientific) and PVPP. The crushed powder was transferred to a chilled Erlenmeyer flask and 200 mL (10 mL/g of stem) of 50 mM Bis-Tris Propane (pH 8.5) containing 10% (v/v) glycerol, 1% (w/v) polyvinylpyrrolidone (PVP-40) and 5 mM DTT was added. The sample was vigorously shaken for 30 s, filtered through one layer of Miracloth (Calbiochem) and centrifuged at 15,000 *x g* for 20 min at 4°C. The soluble protein fraction was transferred to a clean, chilled Erlenmeyer flask and a final 30% (w/v) saturation of ammonium sulfate was added. The sample was then stirred for 5 min and incubated for 1 h at 4°C followed by centrifugation (as above). The precipitated proteins and polyphenols were kept to test for cathinone aminotransferase/synthase activity. Additional ammonium sulfate was added to reach a final 60% (w/v) saturation level followed by stirring for 5 min and incubation for 1 h at 4°C. The sample was centrifuged (as above) and precipitated proteins kept for analysis. Additional ammonium sulfate was again added to reach a final 90% (w/v) saturation level followed by stirring, incubation and centrifugation (as above). Each precipitated protein fraction was dissolved in 2.5 mL of 100 mM HEPES-NaOH (pH 8.0) containing 1 mM EDTA and desalted using PD-10 columns. The proteins were concentrated using Amicon Ultra-4 10,000 MWCO filters and stored at -80°C in 100 mM HEPES-NaOH (pH 8.2) containing 1 mM sodium EDTA

and 20% (v/v) glycerol. Protein samples were evaluated for their protein content using the Bradford assay with BSA as the standard.

**2.8 Enzyme Assays**

Recombinant proteins were assayed for aromatic aminotransferase and cathinone aminotransferase/synthase activities. Enzyme kinetics for the forward aromatic aminotransferase reactions were determined by monitoring the absorbance of reaction products by spectrophotometry as previously described (Collier and Kohlhaw, 1972). Production of 4-hydroxyphenylpyruvate was monitored at 331 nm ($\varepsilon_{331}$ = 19,920 $M^{-1}$ $cm^{-1}$), phenylpyruvate at 320 nm ($\varepsilon_{320}$ = 17,500 $M^{-1}cm^{-1}$) and indole-3-pyruvate at 328 nm ($\varepsilon_{328}$ = 10,000 $M^{-1}cm^{-1}$) corresponding to L-tyrosine, L-phenylalanine and L-tryptophan, respectively. The assay contained 100 mM Tris-HCl buffer (pH 9.0), 0.1 μg recombinant protein (0.01 μg for L-tyrosine), 0.1 mM sodium EDTA and the following substrate concentrations: 0.3 to 6 mM amino donor (L-tyrosine, L-phenylalanine or L-tryptophan) and 0.1 to 1.5 mM amino acceptor (α-ketoglutarate) in 250 μL total volume. Reactions were incubated at 30°C for 30 min and terminated by the addition of 70 μL of 2 N sodium hydroxide. Enzyme assays excluding recombinant enzyme, excluding a substrate or using heat-inactivated enzyme were used as negative controls. Reactions were incubated at room temperature for 30 min then transferred to Greiner UV-Star 96-well, flat bottom plates and scanned by a Bio-Tek PowerWave XS plate reader. Protein extracts from *Ephedra sinica* stem tissue were also assayed for forward tyrosine aminotransferase activity, as described above.

Enzyme kinetics for the reverse aromatic aminotransferase and cathinone aminotransferase/synthase reactions were determined by coupling the reaction with

α-ketoglutarate dehydrogenase (Sigma) and monitoring the absorbance of NADH at 340 nm ($\varepsilon_{340}$ = 6,300 $M^{-1}cm^{-1}$), as previously described (Prabhu and Hudson, 2010). The assay contained 100 mM Tris-HCl buffer (pH 9.0), 2 μg recombinant protein, 0.1 mM sodium EDTA, 600 mg α-ketoglutarate dehydrogenase (0.28 U/mg), 0.3 mM Coenzyme A, 0.3 mM $NAD^+$ and the following substrate concentrations: 0.3 to 25 mM amino donor (L-glutamate) and 0.05 to 0.9 mM 4-hydroxyphenylpyruvate, 0.05 to 0.9 mM phenylpyruvate, 0.01 to 0.3 mM indole-3-pyruvate, 0.01 to 1.0 mM 1-phenylpropane-1,2-dione, 0.01 to 1.0 mM (*S*)-phenylacetylcarbinol in a total volume of 500 μL. (*S*)-phenylacetylcarbinol was purchased from Toronto Research Chemicals as an 80% enantiomeric excess mixture. Reactions were incubated at 30°C for 30 min then immediately transferred to Greiner UV-Star plates (see above) and scanned in the plate reader. Protein extracts from *Ephedra sinica* stem tissue was also assayed for cathinone aminotransferase/synthase activity, as described above. Beer's Law was used to determine concentrations of the reaction products. Apparent $V_{max}$ and $K_m$ values were calculated using Lineweaver-Burk (double reciprocal) plots for analysis of the Michaelis–Menten equation.

**2.8 Multiple Sequence Alignment**

Deduced amino acid sequences for tyrosine aminotransferase (TAT) enzymes from *Arabidopsis thaliana* (GenBank accession number: AED94050.1), *Papaver somniferum* (ADC33123.1), *Homo sapiens* (EAW59231.1) (Sivaraman and Kirsch, 2006) and *Trypanosoma cruzi* (AAA02975.1) (Montemartini *et al*., 1993) as well as an aromatic aminotransferase from *Cucumis melo* (ADC45389.1) were obtained from UniProt (uniprot.org). Multiple sequence alignment was performed using Clustal Omega on the

EBI servers (http://www.ebi.ac.uk/Tools/msa/ clustalo). BoxShade (http://www.ch.embnet.org/soft-ware/BOX_form.html) was used to generate a shaded image based on sequence similarity. Residues shaded black are identical in all aligned sequences; residues in boxes are similar in at least three of the aligned sequences.

# 3. Results

## 3.1 *Ephedra sinica* Plant Growth

Growing *Ephedra sinica*, in a greenhouse none the less proved to be quite difficult. It was determined that a 50% sand, 50% Cactus Soil mixture was most efficient for germinating and growing *Ephedra sinica* (data not shown). The seedlings grew well in the autumn, winter and spring months, sprouting new shoots and producing fresh stems often. However, during humid summer months the plants became dormant and would not produce new stem tissue. Many died or did not recover as the humidity in summer months usually exceeded 80%. Figure 3.1 is a photograph of a one year old *Ephedra sinica* plant that survived the summer months and began producing new stems in early autumn. The dashed arrows designate young stem tissue utilized in this study. This plant was sown in November. Stem tissues were collected from it throughout the winter and spring as they appeared. In late April the plant became dormant and was not harvested until it began producing new stem tissue again in September. This plant was utilized to extract protein. Plants utilized for RNA extraction were grown in the same manner.

## 3.2 RNA Extraction

Total RNA was extracted from young *Ephedra sinica* stem tissue (Figure 3.2) by means of a phenol/chloroform method. The 28S (1.5 Kb) and 18S (1 Kb) ribosomal RNA (rRNA) bands are clearly visible as well as 5S rRNA (850 bp) and possibly transfer RNA (tRNA). There is also very little smearing indicating the RNA is intact and of high quality. This RNA was used as a template for high-throughput next-generation sequencing and to generate cDNA for cloning of candidate genes.

### 3.3 Next-generation Sequencing and Gene Triage

*Ephedra sinica* RNA was sequenced using the Illumina HiSeq 2000 platform at the McGill University and Génome Québec Innovation Centre. As indicated in Table 3.1 this high-throughput sequencing method generated more than 95 million, 108 bp, paired-end reads while sequencing over 19 billion nucleotide bases. This massive collection of data was assembled and annotated by Dr. Christoph Sensen's laboratory at the University of Calgary in collaboration with Genome Canada's PhytoMetaSyn project. The assembly was completed using the Velvet algorithm (Zerbino and Birney, 2008) and placed in the MAGPIE online portal. As shown in Table 3.2 more than 89 million nucleotide bases in 291 thousand reads were assembled into more than 59 thousand contigs, 36 thousand singletons and almost 23 thousand coding DNA sequences (CDSs). This EST database formed the functional genomics platform utilized by all collaborating laboratories to study ephedrine alkaloid biosynthesis in *Ephedra sinica*. This database partially confirmed the proposed biosynthetic pathway by providing a number of candidate genes for each putative enzyme in the pathway.

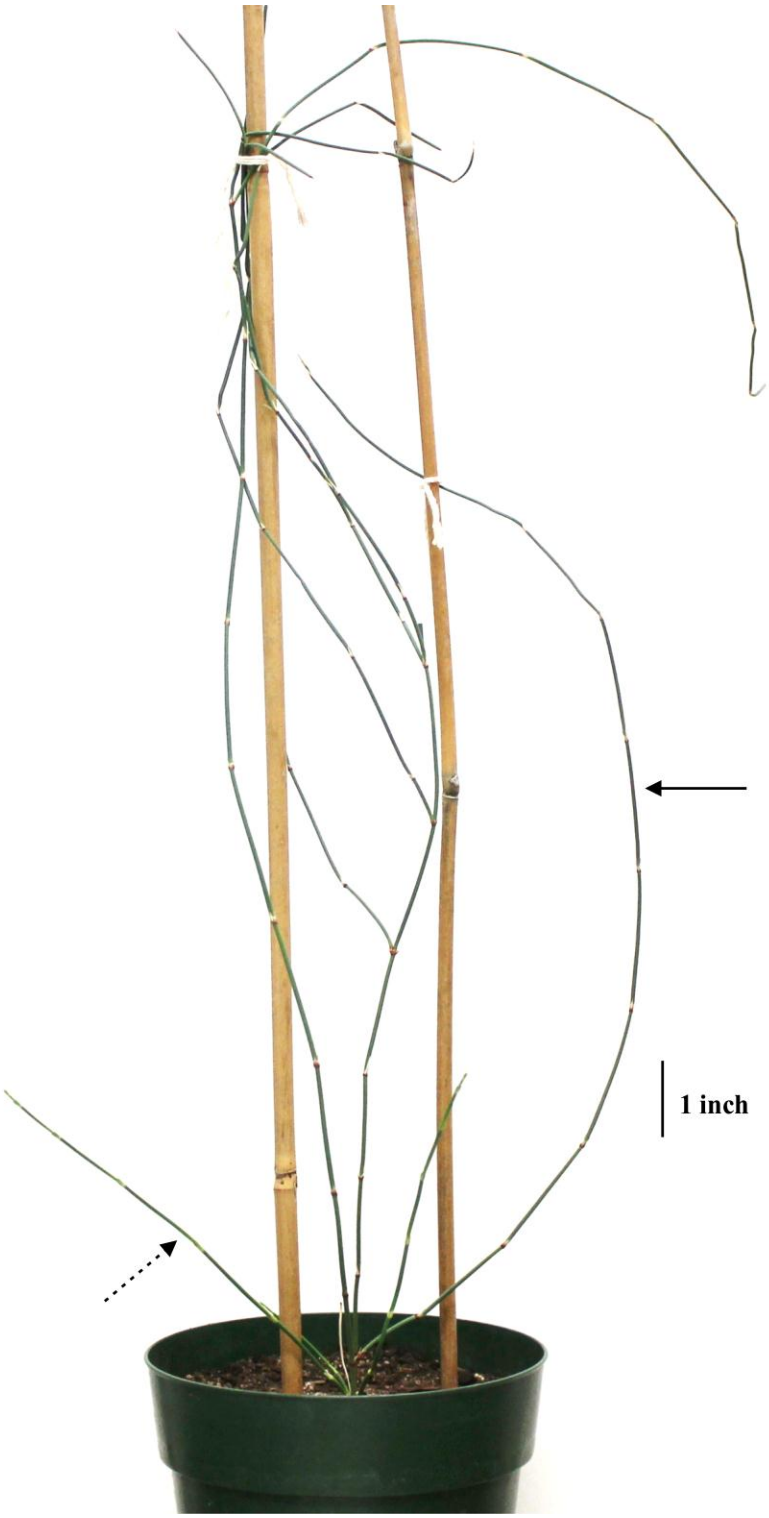


**Figure 3.1 Photo of one year old *Ephedra sinica* plant grown in a greenhouse.** Freshly growing, light green parts of the stem (noted with dashed arrow) are 'young' tissue used in this study. The dark green parts of the stem (solid arrow) are considered 'mature' tissue.

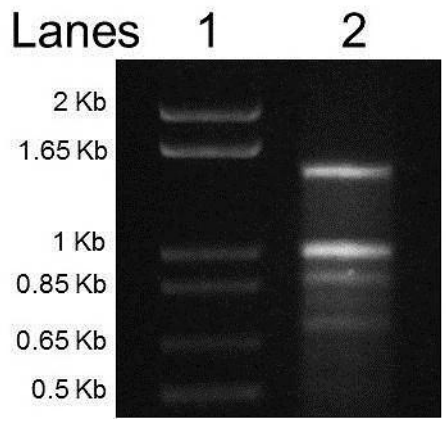


**Figure 3.2 Agarose gel electrophoresis of extracted *Ephedra sinica* RNA.** After the sample was heated to 70°C for 5 minutes, 1 µg of *Ephedra sinica* RNA was run on a 1% (w/v) agarose gel and visualized with RedSafe nucleic acid stain (iNtRON Biotechnology). Lane 1: 1 Kb Plus DNA Ladder (Invitrogen). Lane 2: extracted RNA sample. The 28S (1.5 Kb) and 18S (1 Kb) rRNA bands are clearly visible as well as 5S rRNA (850 bp) and possibly tRNAs.

**Table 3.1 Next-generation sequencing data produced by Illumina HiSeq 2000 on *Ephedra sinica* RNA.**

| Run Type | Application | Read Length | Number of Reads | Number of Bases | Passed Filter |
|---|---|---|---|---|---|
| Paired-end | RNA-Seq | 108 bp | 95,676,077 | 19,135,215,400 | 96% |

**Table 3.2 MAGPIE Velvet assembly of *Ephedra sinica* next-generation sequencing reads.**

| Number of Bases | Number of Contigs | Number of Singletons | Number of CDSs | Number of Reads Assembled |
|---|---|---|---|---|
| 89,621,530 | 59,448 | 36,574 | 22,874 | 291,623 |

### 3.4 Candidate Gene Selection

Candidate PLP-dependent aminotransferases (Table 3.3) were selected based on sequence similarity to previously characterized aromatic aminotransferase enzymes. Other fold I type (aspartate superfamily) aminotransferase enzymes were also used because they constitute the majority of known aminotransferases, including aromatic aminotransferases (Eliot and Kirsch, 2004). The MAGPIE portal allows this database to be searched using BLAST (BLASTn or tBLASTn), a Gene Ontology term, a sequence identifier or a keyword search. The database was first searched by tBLASTn using the amino acid sequences from *Arabidopsis thaliana* (Prabhu and Hudson, 2010) and *Papaver somniferum* (Lee and Facchini, 2011) tyrosine aminotransferase and *Cucumis melo* aromatic aminotransferase (Gonda *et al*., 2010). This revealed a large number of candidates which were expanded when the sequence for *Rosa* aromatic aminotransferase and was published (Hirata *et al*., 2012). To ensure all fold I type aminotransferases were accounted for in the list of candidate genes, the database was also searched using BLAST and a variety of plant aminotransferases (see Table 3.3). Aminotransferase enzymes from humans, fungi and bacteria were also attempted however they did not show the significant sequence similarity (greater than 40%) observed when using plant enzymes as a query.

**Table 3.3 Candidate PLP-dependent transaminases/aminotransferases from *Ephedra sinica* EST collection putatively involved in the formation of (*S*)-cathinone from 1-phenylpropane-1,2-dione precursor**

| tBLASTn Query | | | Top Hits in *Ephedra sinica* EST Collection | | | |
|---|---|---|---|---|---|---|
| ***Accession*** | ***Enzyme*** | ***Organism*** | ***EsUniGene*** | ***Members*** | ***Score/Evalue*** | ***Identities/Positives (%)*** |
| ADM67557.1 | Prephenate AT | *Petunia hybrida* | Contig4103 | 7 | 657/0.0 | 71/88 |
| | | | Contig12559 | 8 | 447/1.00E-137 | 58/76 |
| ADC45389.1 | Aromatic AT | *Cucumis melo* | Contig13244 | 2 | 415/1.00E-116 | 45/70 |
| | | | Singlet18529 | | 387/1.00E-107 | 43/67 |
| | | | Contig12701 | 3 | 373/1.00E-103 | 44/68 |
| ADC33123.1 | Tyrosine AT | *Papaver somniferum* | Contig12701 | 3 | 393/1.00E-109 | 47/69 |
| | | | Singlet18529 | | 391/1.00E-108 | 46/67 |
| | | | Contig13244 | 2 | 389/1.00E-108 | 44/68 |
| AEE35079.1 | Tryptophan AT | *Arabidopsis thaliana* | Contig19156 | 2 | 378/1.00E-104 | 48/65 |
| | | | Contig901 | 10 | 277/3.00E-74 | 40/58 |
| | | | Contig4535 | 4 | 275/2.00E-73 | 39/57 |
| AB669189.1 | Aromatic AT | *Rosa* 'Yves Piaget' | Contig18610 | 15 | 112/1.00E-24 | 27/45 |
| | | | Contig287 | 17 | 91/4.00E-18 | 23/43 |
| | | | Contig16680 | 3 | 81/2.00E-20 | 28/44 |
| | | | Contig29549 | 12 | 71/5.00E-17 | 25/45 |
| CAF31327.1 | Aspartate AT | *Pinus pinaster* | Singlet4072 | | 133/7.00E-31 | 28/48 |
| | | | Contig19895 | 8 | 70/1.00E-11 | 25/42 |
| AAB46610.1 | Aspartate AT | *Medicago sativa* | Contig11412 | 26 | 655/0.0 | 76/86 |
| | | | Contig22464 | 16 | 459/1.00E-129 | 53/73 |
| | | | Contig798 | 16 | 417/1.00E-117 | 51/68 |
| | | | Contig12253 | | *Found using Keyword Search | |

Sequence alignments (at NCBI) were then performed to determine which previously characterized enzyme had the highest sequence similarity with each candidate. Candidates were also aligned with each other. Some candidates were discarded if they had higher than 97% nucleotide sequence similarity with another candidate as they would be impossible to clone independently. These could be alleles of the same gene; a high degree of polymorphisms were observed not only in the EST database but also when comparing the Sanger sequencing results of cloned candidate genes to their corresponding entry in the database. The seeds utilized in this study were collected from wild, open-air pollinated *Ephedra sinica* populations which could explain the possible heterozygosity and variation observed in candidate cDNA sequences.

Contigs 4103 and 12559 were selected due to their high similarity with the *Petunia* prephenate aminotransferase (Maeda *et al*., 2011); Contigs 13244, 12701 and Singlet 18529 were selected based on similarity with *Papaver somniferum* tyrosine aminotransferase (Lee and Facchini, 2011) and *Cucumis melo* aromatic aminotransferase (Gonda *et al.*, 2010); Contigs 19156, 901 and 4535 based on similarity with *Arabidopsis thaliana* tryptophan aminotransferase (Stepanova, Robertson-Hoyt *et al*. 2008); Contigs 18610, 287, 16680 and 29549 based on similarity with *Rosa* 'Yves Piaget' aromatic aminotransferase (Hirata *et al*., 2012); Singlet 4072 and Contig 19859 based on similarity with *Pinus pinaster* aspartate aminotransferase (de la Torre *et al*., 2006); Contigs 11412, 22464 and 798 based on similarity with *Medicago sativa* aspartate aminotransferase (Udvardi and Kahn, 1991). Contig 12253 was found using the keyword search 'aminotransferase' on MAGPIE.

After selecting candidate genes, the longest deduced amino acid sequence for each was determined and PCR primers were designed (Appendix A1). Candidate genes were amplified by PCR from *Ephedra sinica* cDNA and cloned into their respective protein expression vectors (see Chapter 2.4).

### 3.5 Heterologous Expression and Purification of Recombinant Proteins

Contigs 13244 and 12701 were cloned into the pQE30 expression vector and expressed in *E. coli* strain XL10-Gold. Figure 3.3 presents a PAGE gel of purified recombinant 13244 protein (Lane 4). Figure 3.4 presents a PAGE gel of purified recombinant 12701 (Lane 3). Purified protein extracts from this expression system were quite pure as very few bands representing undesired proteins were observed. Expression of 4103 and 12559 cDNA was also attempted in this system; however recombinant protein was not detectable in soluble or crude protein extracts. The gene was properly inserted into the expression vector as confirmed by restriction digest and Sanger sequencing. Expression of 4103 and 12559 cDNA was then attempted in the pPROEX HTb vector because it reduces basal levels of expression by requiring T7 RNA polymerase. The recombinant protein may have been toxic to the *E. coli*.

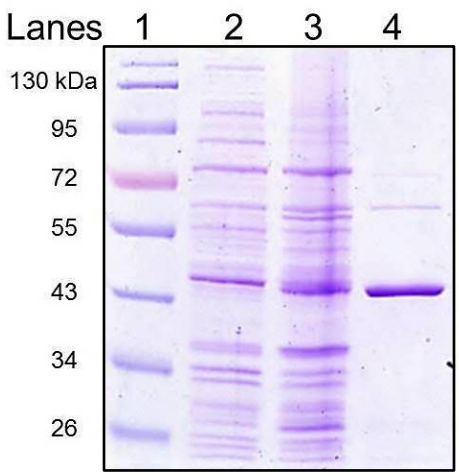


**Figure 3.3 Heterologous expression and purification of *Ephedra sinica* Contig13244 in *E. coli* strain XL10-Gold.** Protein extracts were obtained from *E. coli* transformed with the pQE30+Contig13244 construct. Lane 1: PageRuler Plus (Fermentas) – molecular mass marker (kDa), lanes 2 and 3: crude protein extracts from uninduced and induced cells, respectively. Lane 4: 1.0 µg of purified recombinant enzyme. The protein samples were resolved on a 10% (w/v) SDS-PAGE gel and visualized with Coomassie R-250 stain (Bio-Rad).

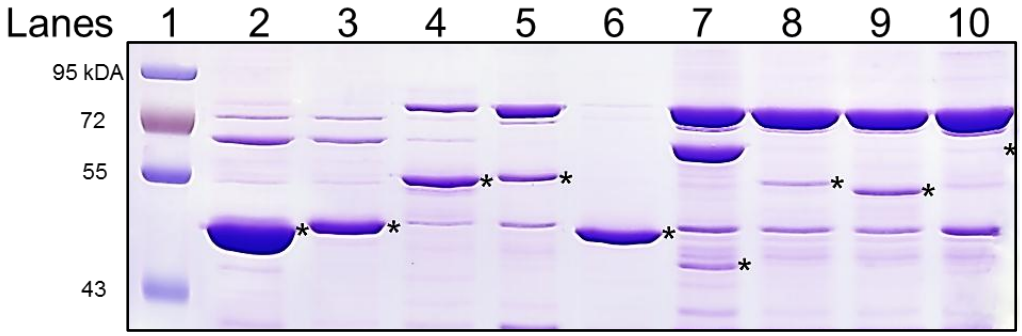


**Figure 3.4 Heterologous expression and purification of various *Ephedra sinica* aromatic aminotransferase candidates in *E. coli*.** Protein extracts were obtained from *E. coli* transformed with the following constructs: pQE30+ Contig13244 (Lane 2) and pQE30+Contig12701 (Lane 3) expressed in *E. coli* strain XL10-Gold, pProEX+Contig4103 expressed in *E. coli* OverExpress strain (Lane 4), Contig 287, Singlet 4072, Contigs 798, 19895, 22464 and 12253 (Lanes 5 – 10, respectively) were cloned into pLATE51 vectors and expressed in *E. coli* strain BL21 (DE3). Lane 1: PageRuler Plus – molecular mass marker (kDa). The protein samples were resolved on a 10% (w/v) SDS-PAGE gel and visualized with Coomassie R-250 stain (Bio-Rad). Asterisks (*) identify the recombinant proteins of interest in each sample, based on their predicted molecular mass.

Contig 4103 (Figure 3.4; Lane 4) was successfully expressed using the pPROEX HTb expression vector and in *E. coli* strain OverExpress. Contig 12559 was cloned into the same expression system; however the recombinant protein was not present in soluble or crude protein extracts. The gene insert was properly cloned in-frame as confirmed by restriction digest and Sanger sequencing. Contig 12559 as well as the remaining aminotransferase candidates were cloned into the pLATE51 LIC expression vector (requires T7 RNA polymerase as well). The LIC system makes cloning much more efficient so it was chosen for the remaining cDNA candidates.

Figure 3.5 presents a PAGE gel of purified recombinant 12559 protein and is a typical gel for expression of candidate genes in the pLATE51 expression system. Constructs in the pLATE51 LIC system were expressed in *E. coli* strain BL21 (DE3). There is no noticeable difference between the uninduced and induced crude protein controls (Figure 3.5; Lanes 2 and 3, respectively). The protein of interest was expected to be more visible in the induced control compared to the uninduced. Also, there is a considerable amount of undesired proteins in the purified sample (Figure 3.5; Lane 5). Recombinant 287, 4072, 798, 19895, 22464 and 12253 proteins were all successfully purified in this system as well (Figure 3.4; Lanes 5 – 10, respectively). As with recombinant 12559 protein, all except 4072 had considerable amounts of undesired proteins present in the purified samples. For this reason, larger amounts of these protein samples (5 – 10 µg) were used for assaying their enzymatic functions.

Recombinant protein 18610 was successfully expressed as a GST fusion protein (Figure 3.6; Lane 4). Again, very little difference is observed between the uninduced and induced controls (Figure 3.6; Lanes 2 and 3, respectively). Few undesired protein bands are observed since these recombinant proteins were cleaved from the GST fusion during purification.

Unfortunately, the remaining six candidate cDNAs – Contigs 19156, 901, 4535, 16680, 29549 and 11412 were not successfully expressed in any of the systems attempted. Singlet 18529 could not be amplified from the cDNA library, possibly because it is not well represented (singleton).

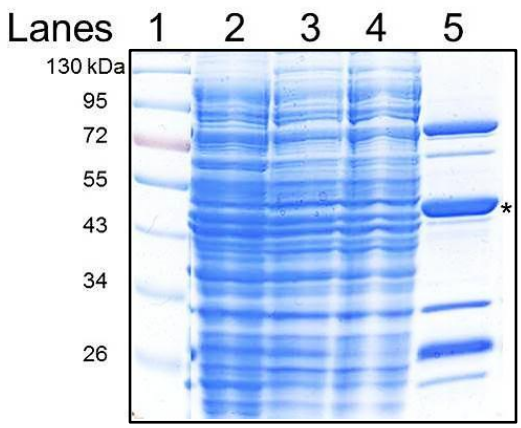


**Figure 3.5 Heterologous expression and purification of *Ephedra sinica* Contig12559 in *E. coli* strain BL21 (DE3).** Protein extracts were obtained from *E. coli* transformed with the pLATE51+Contig12559 construct. Lane 1: PageRuler Plus – molecular mass marker (kDa), lanes 2 and 3: crude protein extracts from uninduced and induced cells, respectively. Lane 4: soluble protein extract. Lane 5: 3.0 µg of purified recombinant enzyme. The protein samples were resolved on a 10% (w/v) SDS-PAGE gel and visualized with Coomassie G-250 stain (Bio-Rad). An asterisk (*) identifies the recombinant protein of interest, based on its predicted molecular mass.

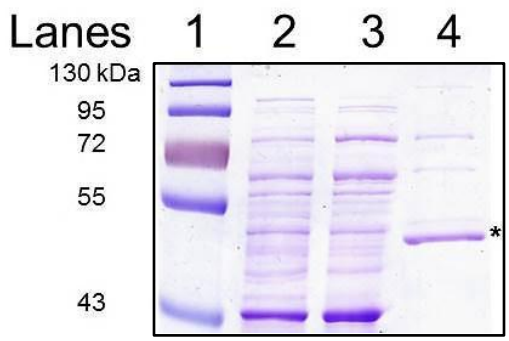


**Figure 3.6 Heterologous expression and purification of *Ephedra sinica* Contig 18610 in *E. coli* strain BL21 (DE3).** Protein extracts were obtained from *E. coli* transformed with the pGEX-6P-2+Contig18610 construct. Lane 1: PageRuler Plus – molecular mass marker (kDa), lanes 2 and 3: crude protein extracts from uninduced and induced cells, respectively. Lane 4: cleaved and purified protein extract. The protein samples were resolved on a 10% (w/v) SDS-PAGE gel and visualized with Coomassie R-250 stain (Bio-Rad). An asterisk (*) identifies the recombinant protein of interest, based on its predicted molecular mass.

## 3.6 Enzyme Assays

### *3.6.1 Recombinant Protein Enzyme Assays*

Recombinant proteins of Contigs 13244 and 12701 were able to successfully catalyze all three aromatic aminotransferase reactions – L-tyrosine to 4-hydroxyphenylpyruvate, L-phenylalanine to phenylpyruvate and L-tryptophan to indole-3-pyruvate – in both the forward and reverse directions. Initial forward L-tyrosine activity detected for recombinant 13244 and 12701 was 0.462 and 0.169 μmoles per min per mg protein, respectively (n=3). Interestingly, only recombinant 12701 required exogenous PLP in enzyme assays to function, while recombinant 13244 did not. Since recombinant 13244 had 2.7-fold greater activity than 12701 under identical conditions, it was chosen for full characterization.

After optimizing the temperature (30°C), time (30 min) and amount of recombinant 13244 enzyme used (0.01 – 0.1 μg) the optimum pH was determined for the forward L-tyrosine reaction (Figure 3.7). Tris-HCl (also known as Trizma) and TAPS buffers were used for pH 8.0 – 9.0, CHES for pH 9.0 – 10.0 and CAPS for pH 10.0 – 11.0. There appeared to be two peaks, one at pH 9.0 and the other at pH 10.0. Since the highest activity was observed with Tris-HCl buffer at pH 9.0, it was chosen as the buffer for all enzyme assays. All enzyme assays were repeated at least three times with similar results.

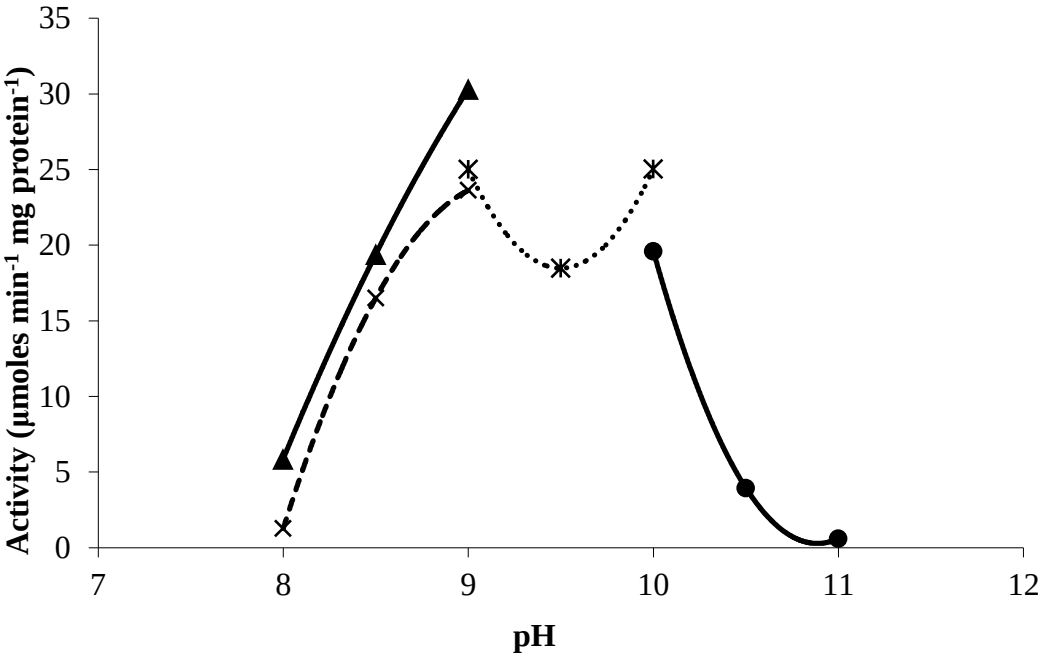


**Figure 3.7 pH Optima of purified recombinant *Ephedra sinica* Contig13244.** Forward L-tyrosine aminotransferase activity was determined from the increase in absorbance at 331 nm, corresponding to the conjugate oxo-acid 4-hydroxyphenylpyruvate. Assays were performed for 30 min at 30°C in 100 mM of each buffer (Trizma/Tris-HCl, TAPS, CHES and CAPS) at the indicated pH values in the presence of 0.05 µg of purified enzyme, 3 mM L-tyrosine, 1 mM α-ketoglutarate and 0.1 mM EDTA. The molar extinction coefficient of 4-hydroxyphenylpyruvate was used along with Beer's Law to calculate the quantity of each reaction product. Values represent the mean specific activity measured in µmoles per min per mg protein (n=3). Tris-HCl (also known as Trizma) buffer at pH 9.0 was chosen for characterization as it produced the highest activity. Triangles represent Trizma/Tris-HCl buffer, X represents TAPS, X with vertical lines represent CHES and circles represent CAPS buffer.

Enzymatic activities for the three forward aromatic aminotransferase reactions are compared in Figure 3.8. Recombinant 13244 preferred to utilize L-tyrosine as a substrate in the forward direction. Activities for the three reverse reactions are compared in Figure 3.9. In the reverse direction, 4-hydroxyphenylpyruvate and phenylpyruvate were the preferred substrates. Apparent kinetic parameters were determined using Lineweaver-Burk (double-reciprocal) plots to analyze the Michaelis–Menten equation. Figure 3.10 presents the triplicate plots used to calculate $V_{max}$ and $K_m$ values for the forward L-phenylalanine reaction as an example. Tables 3.4a and 3.4b present the kinetic parameters for the forward and reverse aromatic aminotransferase reactions, respectively. In the forward direction, the reaction was most efficient using L-tyrosine as indicated by a $k_{cat}/K_m$ value which was significantly greater according to Fisher's protected least significant difference (LSD) test ($p \leq 0.05$). In the reverse direction, the highest reaction rate ($V_{max}$) and turnover rate ($k_{cat}$) was achieved with phenylpyruvate. Hydroxyphenylpyruvate had the lowest $K_m$ value and highest efficiency ($k_{cat}/K_m$), significantly different according to Fisher's protected LSD test ($p \leq 0.05$). Unfortunately, none of the additional candidate cDNAs showed any detectable aminotransferase activity towards L-tyrosine, 1-phenylpropane-1,2-dione or (*S*)-phenylacetylcarbinol.

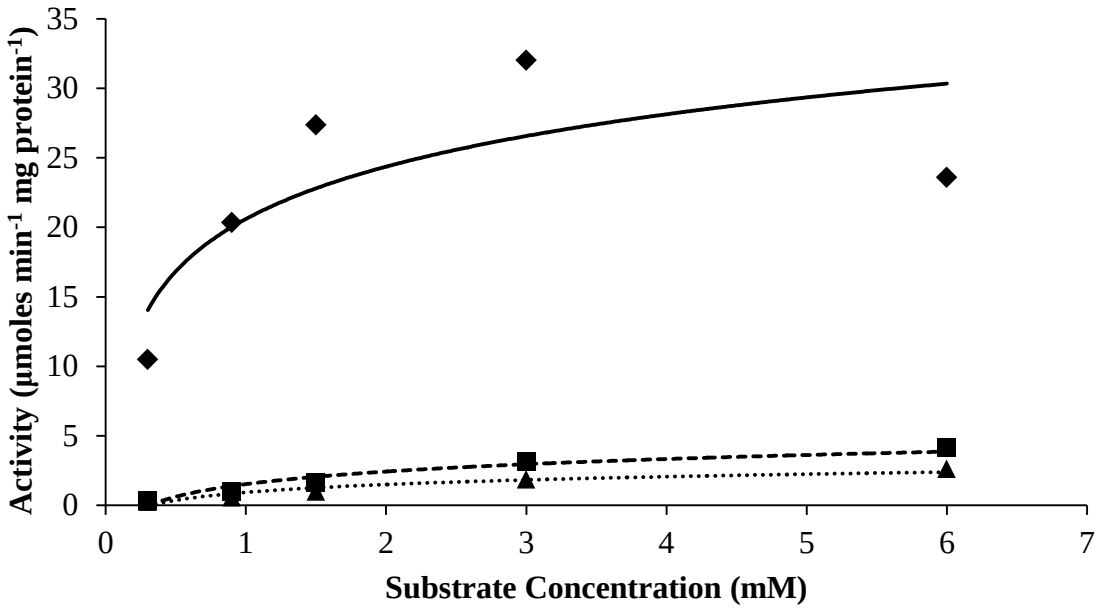


**Figure 3.8 Steady-state enzyme kinetics of the forward reactions of purified recombinant *Ephedra sinica* Contig13244.** $V_{max}$ and $K_m$ values were calculated using Lineweaver-Burk (double reciprocal) plots for analysis of the Michaelis–Menten equation. Different concentrations of L-tyrosine, L-phenylalanine and L-tryptophan were used as amino donors. Enzyme activity was determined from the increase in absorbance at 331, 320 and 328 nm corresponding to the conjugate oxo-acids 4-hydroxyphenylpyruvate, phenylpyruvate and indole-3-pyruvate, respectively. Molar extinction coefficients were used along with Beer's Law to calculate the quantity of each reaction product. Incubation time and enzyme concentration were optimized prior to kinetic parameter analysis. Values represent the mean specific activity measured in µmoles per min per mg protein (n=3). Diamonds represent L-tyrosine, squares represent L-phenylalanine and triangles represent L-tryptophan.

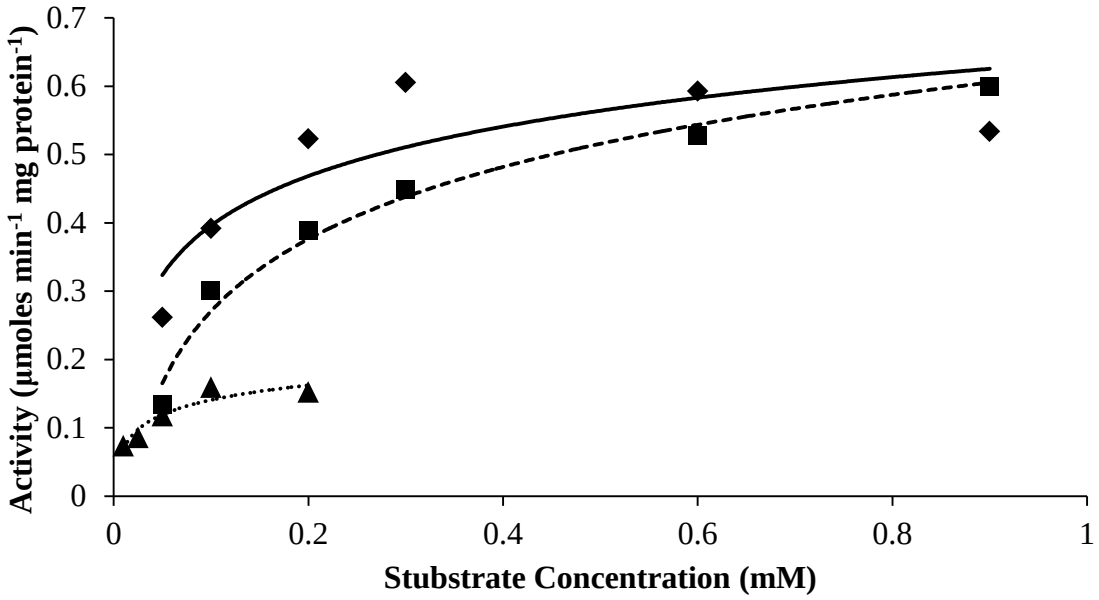


**Figure 3.9 Steady-state enzyme kinetics of the reverse reactions of purified recombinant *Ephedra sinica* Contig13244.** $V_{max}$ and $K_m$ values were calculated using Lineweaver-Burk (double reciprocal) plots for analysis of the Michaelis–Menten equation. Different concentrations of 4-hydroxyphenylpyruvate, phenylpyruvate and indole-3-pyruvate were used as amino acceptors. The reactions used L-glutamate as amino donor, producing α-ketoglutarate after transamination. Reactions were coupled with α-ketoglutarate dehydrogenase, which produces NADH as a by-product, allowing activity to be determined from the increase in absorbance at 340 nm. The molar extinction coefficient was used along with Beer's Law to calculate the quantity of NADH produced in each reaction. Incubation time and enzyme concentration were optimized prior to kinetic parameter analysis. Values represent the mean specific activity measured in µmoles per min per mg protein (n=3). Diamonds represent 4-hydroxyphenylpyruvate, squares represent phenylpyruvate and triangles represent indole-3-pyruvate.

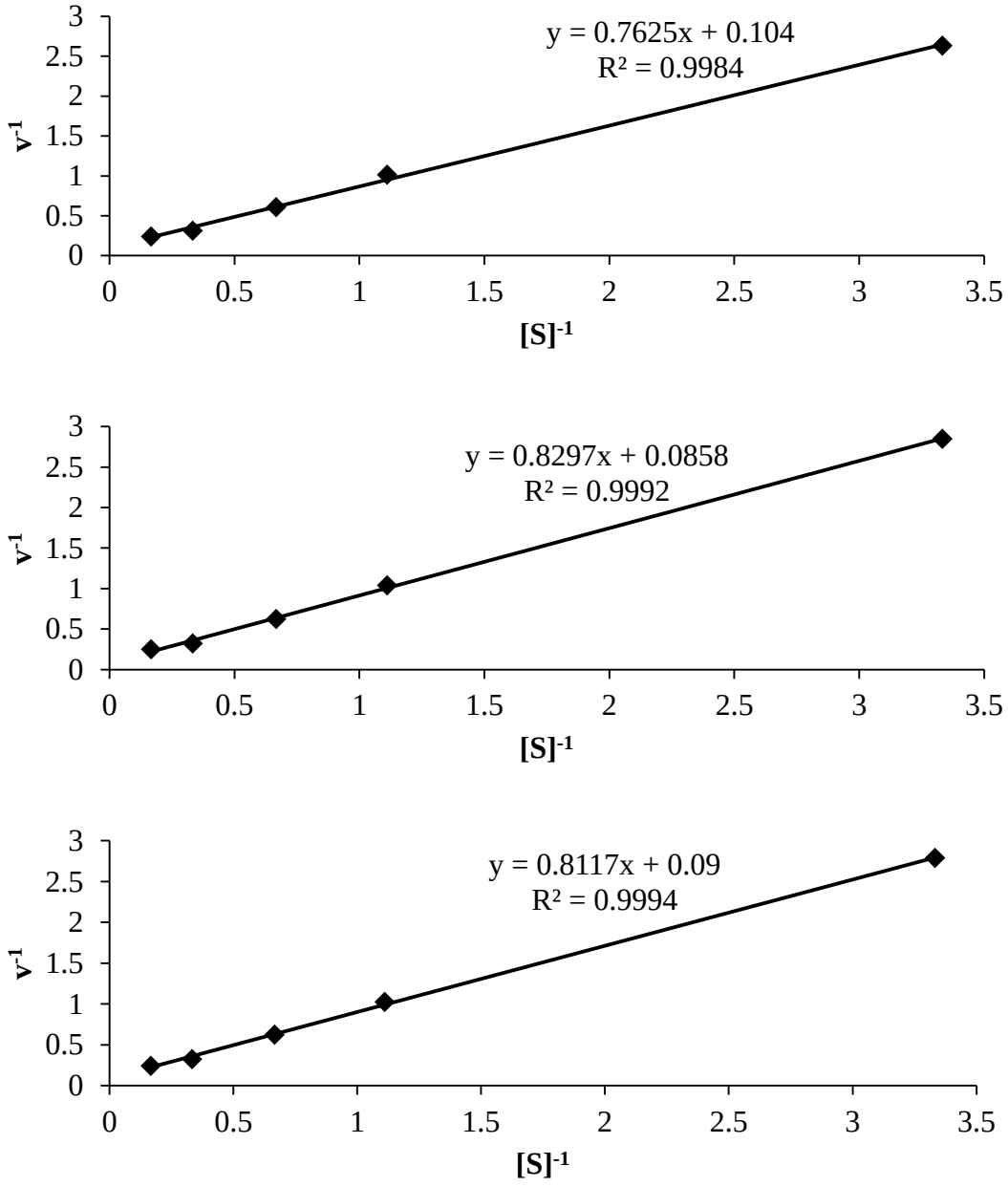


**Figure 3.10 Triplicate Lineweaver-Burk (double reciprocal) plots used for the calculation of kinetic parameters for the forward L-phenylalanine reaction of *Ephedra sinica* Contig13244.** Assays were performed for 30 min at 30°C in 100 mM Tris-HCl buffer (pH 9) in the presence of 0.1 µg purified recombinant protein, varying concentrations of the amino donor L-Phenylalanine, 1.0 mM α-ketoglutarate and 0.1 mM EDTA. The plots display the reciprocals of enzyme activity ($v^{-1}$) and L-phenylalanine concentration ($[S]^{-1}$). $V_{max}$ values were calculated by dividing one by the y-intercept value. $K_m$ values were calculated by dividing one by the x-intercept value.

**Table 3.4a Kinetic parameters of forward Aromatic Aminotransferase reaction for Contig13244 from *Ephedra sinica***

| Substrate | Cosubstrate | $K_m$ (mM) | $V_{max}$ (μmol $min^{-1}$ $mg^{-1}$) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}$ $s^{-1}$) |
|---|---|---|---|---|---|
| L-Tyrosine | α-Ketoglutarate (1 mM) | $0.910 \pm 0.042^{c}$ | $42.2 \pm 1.07^{a}$ | $32.8 \pm 0.831^{a}$ | $36.1 \pm 0.756^{a}$ |
| L-Phenylalanine | α-Ketoglutarate (1 mM) | $8.67 \pm 1.21^{a}$ | $10.8 \pm 1.06^{b}$ | $8.39 \pm 0.821^{b}$ | $0.972 \pm 0.043^{b}$ |
| L-Tryptophan | α-Ketoglutarate (1 mM) | $5.37 \pm 1.04^{b}$ | $4.54 \pm 0.224^{c}$ | $3.53 \pm 0.174^{c}$ | $0.672 \pm 0.111^{b}$ |
| **Least Significant Difference** | | 1.84 | 1.75 | 1.36 | 0.884 |
| α-Ketoglutarate | L-Tyrosine (3.3 mM) | 0.801 ± 0.0418 | 42.0 ± 2.19 | 32.6 ± 1.71 | 40.8 ± 0.099 |

Note: Means for each parameter ($K_m$, $V_{max}$, $k_{cat}$ and $k_{cat}$ / $K_m$) with the same letter are not significantly different at $p \leq 0.05$ according to Fisher's protected least significant difference (LSD) test (n=3).

**Table 3.4b Kinetic parameters of reverse Aromatic Aminotransferase reaction for Contig13244 from *Ephedra sinica***

| Substrate | Cosubstrate | $K_m$ (mM) | $V_{max}$ (μmol $min^{-1}$ $mg^{-1}$) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}$ $s^{-1}$) |
|---|---|---|---|---|---|
| 4-Hydroxyphenylpyruvate | L-Glutamate (10 mM) | $0.105 \pm 0.018^{b}$ | $0.806 \pm 0.05^{b}$ | $0.627 \pm 0.039^{b}$ | $6.06 \pm 0.930^{a}$ |
| Phenylpyruvate | L-Glutamate (10 mM) | $0.364 \pm 0.040^{a}$ | $1.15 \pm 0.037^{a}$ | $0.894 \pm 0.029^{a}$ | $2.47 \pm 0.202^{c}$ |
| Indole-3-pyruvate | L-Glutamate (10 mM) | $0.037 \pm 0.009^{c}$ | $0.214 \pm 0.038^{c}$ | $0.166 \pm 0.029^{c}$ | $4.51 \pm 0.432^{b}$ |
| **Least Significant Difference** | | 0.051 | 0.086 | 0.109 | 1.21 |
| L-Glutamate | Phenylpyruvate (1 mM) | 9.03 ± 0.613 | 0.937 ± 0.075 | 0.729 ± 0.058 | 0.081 ± 0.001 |

Note: Means for each parameter ($K_m$, $V_{max}$, $k_{cat}$ and $k_{cat}$ / $K_m$) with the same letter are not significantly different at $p \leq 0.05$ according to Fisher's protected LSD test (n=3).

*3.6.2 Extracted Protein Enzyme Assays*

Total protein extracts from *Ephedra sinica* stem tissue were fractionated by using ammonium sulfate. Precipitated protein fractions of 0-30%, 30-60% and 60-90% (w/v) ammonium sulfate were produced and screened for enzymatic activity. Each fraction was screened for forward tyrosine aminotransferase activity as described in Chapter 2.8. The 0-30% (w/v) ammonium sulfate fraction contained precipitated polyphenols and pigments which caused a large amount of interference with the absorbance reading. Therefore, activity could not be screened by spectrophotometry. The 30-60% (w/v) fraction showed the highest level of forward tyrosine aminotransferase activity at 0.863 nmoles per min per mg of total protein extract. The 60-90% (w/v) fraction also exhibited forward tyrosine aminotransferase activity at 0.584 nmoles per min per mg of total protein extract. However, when attempting to screen for cathinone aminotransferase/synthase activity using spectrophotometry, the addition of the coupling enzyme raised the $A_{340}$ of the sample above 1.5. Therefore, CAT activity in plant extracted protein samples could not be screened by spectrophotometry.

**3.7 Multiple Sequence Alignment**

A multiple amino acid sequence alignment comparing *Ephedra sinica* Contigs 13244 and 12701 with aromatic aminotransferase enzymes from *Arabidopsis*, musk melon, opium poppy, human and *Trypanosoma* is presented in Figure 3.11. The conserved lysine residue where PLP binds (Prabhu and Hudson, 2010) is noted with an asterisk (*). The sequences do not show an extremely high degree of similarity with one another as shown in Table 3.5 which compares the amino acid sequence of Contig 13244

with a number of previously characterized aromatic aminotransferase enzymes. The highest similarity was with Rose aromatic/phenylalanine aminotransferase as 47% of the amino acids were exact matches. Positives account (similarity value) for exact matches as well as conservative substitutions (similar amino acids) between the two sequences. Tyrosine aminotransferase from *Arabidopsis* and aromatic aminotransferase from melon each had 70% similar sequences with Contig 13244. As observed in *Arabidopsis* (Prabhu and Hudson, 2010), there is often less than 50% identity between aromatic aminotransferases when comparing enzymes from the same species or from different plant species. The same is seen within *Ephedra sinica* when comparing Contigs 13244 and 12701, and when comparing Contig 13244 to previously characterized plant aromatic aminotransferases (Table 3.5).

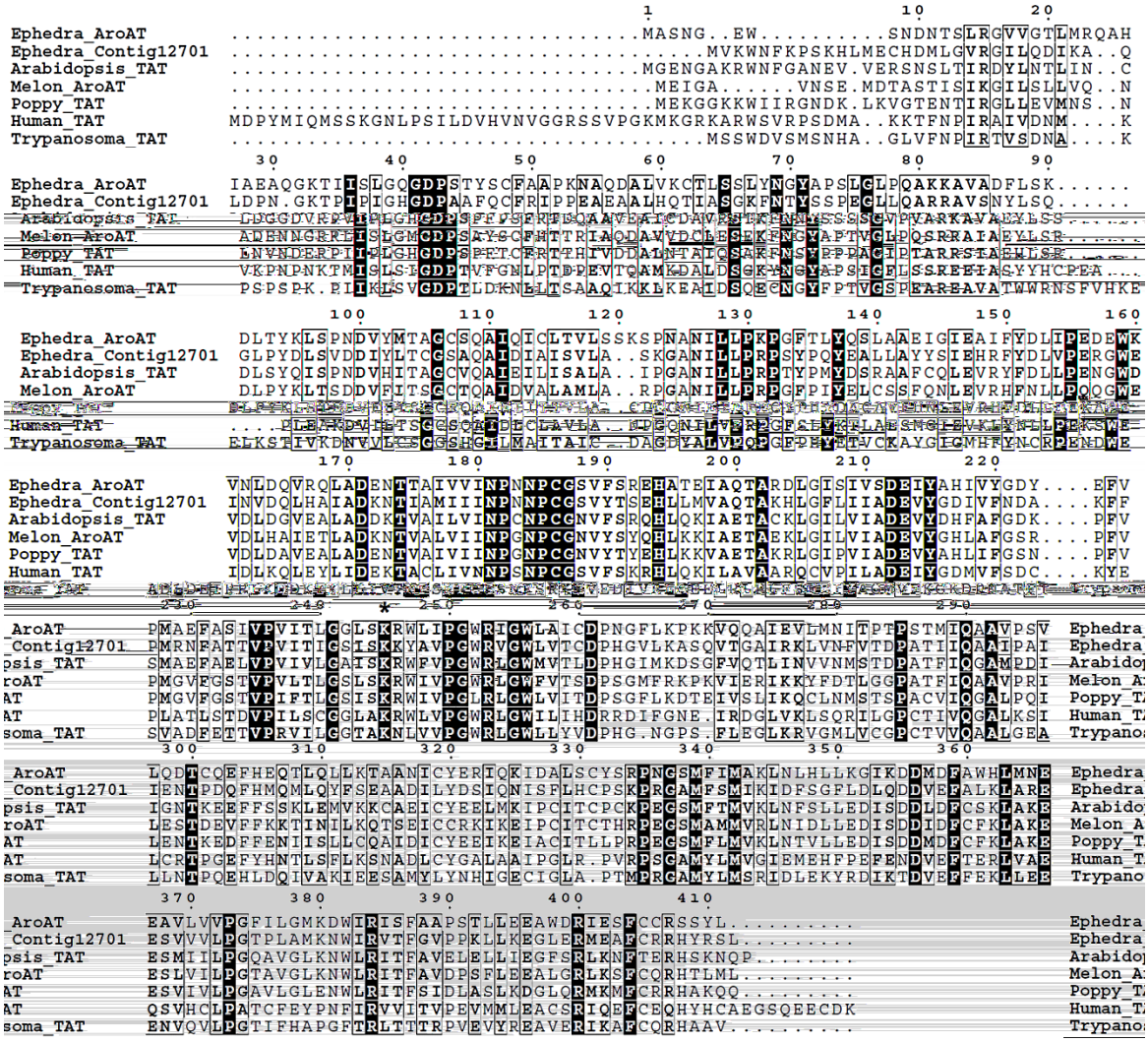


**Figure 3.11 Multiple sequence alignment of tyrosine and aromatic aminotransferase enzymes.** Deduced amino acid sequences for *Ephedra sinica* Contig13244 (AroAT) and Contig12701 were aligned with tyrosine aminotransferases (TAT) from *Arabidopsis thaliana* (GenBank accession number: AED94050.1), *Papaver somniferum* (ADC33123.1), *Homo sapiens* (EAW59231.1) and *Trypanosoma cruzi* (AAA02975.1) as well as an aromatic aminotransferase from Musk Melon – *Cucumis melo* (ADC45389.1). Clustal Omega (EBI) was used to generate the sequence alignment and BOXSHADE was used to shade the alignment based on similarity. Amino acids shaded black are identical in all aligned sequences; residues in boxes are similar in greater than 40% of aligned sequences. The conserved lysine residue where PLP binds is denoted with an asterisk (*).

**Table 3.5 *Ephedra sinica* Contig 13244 sequence similarity with previously characterized aromatic aminotransferase enzymes from various organisms.**

| *Accession* | *Function* | *Organism* | *Score/Evalue* | *Identities/Positives (%)* |
|---|---|---|---|---|
| Contig12701* | Aromatic Aminotransferase | *Ephedra sinica* | 377/1.00E-131 | 46/68 |
| AB669189.1 | Phenylalanine Aminotransferase | *Rosa 'Yves Piaget'* | 379/4.00E-132 | 47/67 |
| AED94050.1 | Tyrosine Aminotransferase | *Arabidopsis thaliana* | 383/1.00e-106 | 46/68 |
| AED96434.1 | Tyrosine Aminotransferase | *Arabidopsis thaliana* | 406/6.00E-144 | 46/70 |
| ADC45389.1 | Aromatic Aminotransferase | *Cucumis melo* | 415/1.00E-116 | 45/70 |
| ADC33123.1 | Tyrosine Aminotransferase | *Papaver somniferum* | 389/1.00E-108 | 44/68 |
| AB182275.1 | Nicotianamine Transaminase | *Oryza sativa Japonica Group* | 396/1.00E-110 | 44/69 |
| EAW59231.1 | Tyrosine Aminotransferase | *Homo sapiens* | 279/4.00E-93 | 36/58 |
| AAA02975.1 | Tyrosine Aminotransferase | *Trypanosoma cruzi* | 223/9.00E-72 | 33/53 |

# Chapter 4. Discussion

## 4.1 Functional Genomics Platform Development

Next-generation sequencing data generated in this study was utilized to develop an annotated EST database, an invaluable tool that provides the foundation for further functional genomic research on *Ephedra* and other ephedrine alkaloid producing species. For the first time, next-generation sequencing has been performed on an ephedrine alkaloid producing plant species in order to elucidate biosynthesis of these pharmaceutically important alkaloids. A snap-shot of the young *Ephedra sinica* stem transcriptome was developed by sequencing its cDNA on the Illumina HiSeq 2000 platform. A total of 291,623 reads were assembled into 59,448 contigs and 36,574 singletons using the Velvet algorithm (Zerbino and Birney, 2008). This massive amount of data was assembled and annotated by Dr. Christoph Sensen's laboratory at the University of Calgary. This served as the first functional genomics platform utilized by three collaborating laboratories – Dr. Frédéric Marsolais' at Agriculture and Agri-Food Canada in London, Dr. Peter Facchini's at University of Calgary and Dr. Efraim Lewinsohn's at Newe Yaar Research Center in Israel – to investigate ephedrine alkaloid biosynthesis.

Recently, similar techniques have been used to elucidate secondary metabolite production in other plant species. A transcriptome of *Papaver somniferum* cell cultures was generated by 454 GS-FLX Titanium pyrosequencing to investigate biosynthesis of sanguinarine – a benzylisoquinoline alkaloid (Desgagne-Penix *et al*., 2010). The 454 GS-FLX platform was also used to generate an annotated EST database of *Panax ginseng* to

study ginsenoside biosynthesis (Chen *et al*., 2011). Similar methodologies are currently being applied to 75 plant species that produce commercially important natural products as part of Genome Canada's PhytoMetaSyn Project (http://www.phytometasyn.ca/). These annotated EST databases provide excellent resources for gene discovery and may aid in the full biochemical characterization of the biosynthesis of various high-value secondary metabolites. In this study, the annotated EST database was used to select candidate aromatic aminotransferase-like genes.

### 4.2 Aromatic Aminotransferases Candidates

Candidate biosynthetic enzymes were selected based on sequence similarity to previously characterized fold I type aminotransferases (Table 3.3). Eighteen candidate cDNAs were chosen from the annotated EST database. One, Singlet18529, could not be amplified from cDNA generated from young *Ephedra sinica* stem RNA. This may be due to its low copy number, as it was only represented once in the assembly, or to an erroneous sequence since singletons may represent artifacts, contaminants or fragments of transcripts (Meyer *et al*., 2009). The remaining 17 candidate cDNAs were successfully amplified by PCR and cloned into expression vectors. Heterologous protein expression of candidate aromatic aminotransferase cDNAs in *E. coli* proved to be extremely difficult.

The first expression system was the pQE30 vector in *E. coli* strain XL10-Gold and only 50% of candidate cDNAs attempted in that system were successfully expressed (Contigs 13244 and 12701). Next, pPROEX HTb vector in *E. coli* strain OverExpress was adopted. The system reduces basal expression of the target cDNA because its expression also requires bacteriophage T7 RNA polymerase (Tabor, 2001). Again only 50% of

candidate cDNAs attempted in the system were successful (Contig 4103). The LIC expression system using the pLATE51 vector and *E. coli* strain BL21 (DE3) was utilized for the majority of candidates to decrease the time required for cloning. Since the inserts and vectors do not require restriction enzyme digest and DNA ligation, one can save up to two days of time using a LIC system. Once again only 50% of candidate cDNAs (seven out of 14) attempted in this system were successfully expressed (Contigs 12559, 287, 798, 19895, 22464 and 12553, and Singlet 4072). Finally the GST-fusion system was attempted with the pGEX-6P-2 vector and *E. coli* strain BL21 (DE3). The only commonality between the remaining candidates was their mid to high predicted pI values – all greater than 6.2 and one (Contig 19156) as high as 9.0. Fusion with GST reduced the predicted pI values, however it did not aid in the successful expression of most of the candidate cDNAs. Only one candidate, Contig 18610 was successfully purified using the GST-fusion system.

Numerous methods of induction were also attempted in order to improve expression of candidate cDNAs in each system. Different concentrations (0.5 mM and 1.0 mM) of IPTG were used and were added at various $OD_{600}$ absorbance values (0.4, 0.5, 0.6 and 0.8) corresponding to increasing cell density. The length (4 hours, 6 hours, 12, hours, 18 hours, 24 hours) and temperature (16°C, room temperature, 37°C) of induction were also varied which did result in the successful expression of 11 of the 17 candidate aminotransferase cDNAs. Recombinant proteins were then purified by affinity chromatography and screened for enzymatic activity.

In order to successfully express and purify the remaining candidates, cell-free *in vitro* expression systems should be considered. A number of *in vitro* protein expression

systems are currently available; Thermo Scientific offers a HeLa cell lysate-based expression system that can yield up to 750 μg/mL per reaction (ThermoScientific, 2011); Life Technologies offers a cell-free *E. coli*-based expression system that can yield mg quantities of purified protein (Invitrogen, 2010); New England BioLabs, Qiagen and Sigma-Adrich (NEB, 2012; Qiagen, 2008; Sigma, 2012) also have *in vitro* protein synthesis kits available. Alternative expression options may allow the remaining candidate cDNAs to be expressed and characterized. *In vitro* protein synthesis has recently been used to successfully express and functionally characterize cytotoxic and membrane bound proteins (Avenaud *et al*., 2004; Klammt *et al*., 2006).

### 4.3 Enzymatic Screening of Candidate Aminotransferases

Recombinant proteins successfully expressed from candidate cDNAs were screened for forward tyrosine aminotransferase activity as well as cathinone aminotransferase/synthase activity using 1-phenylpropane-1,2-dione and (*S*)-phenylacetylcarbinol as substrates. Forward tyrosine aminotransferase activity was first detected with recombinant 13244 protein. This enzyme was fully functional without the addition of exogenous PLP as it was able to complete all three forward and reverse aromatic aminotransferase reactions. Purified recombinant 13244 is yellowish-green in colour indicating PLP is already bound to the enzyme (Prabhu and Hudson, 2010). The addition of PLP did not increase its activity. Forward tyrosine aminotransferase activity was also detected with recombinant 12701 protein. This enzyme was able to complete all three reversible aromatic aminotransferase reactions, however it required exogenous PLP and purified recombinant 12701 was colourless. Activity was not detected when PLP was excluded from the enzyme assay. These enzymes were expressed and purified using

identical methodologies so one would expect the cofactor to be bound to both or neither. A possible explanation for this observation can be seen when comparing the sequences of multiple aromatic aminotransferases (Figure 4.1). Exogenous PLP was not necessary for recombinant 13244, *Arabidopsis* (Prabhu and Hudson, 2010), or opium poppy (Lee and Facchini, 2011) aromatic/tyrosine aminotransferases to function. There appears to be two non-synonymous substitutions in the active site of recombinant 12701 where PLP binds. One replaces an arginine residue with a lysine and the other a tryptophan with a tyrosine residue. These mutations may be responsible for PLP not being bound to purified, recombinant 12701 as these changes are not present in the enzymes where PLP is bound after purification. They may also be responsible for the lower level of aromatic aminotransferase activity; Recombinant 13244 had 2.7-fold greater activity then 12701, under identical conditions, so it was chosen for full characterization.

### 4.4 *Ephedra sinica* Aromatic Aminotransferase – Contig13244

Based on the current study, candidate cDNA Contig13244 is the aromatic aminotransferase for *Ephedra sinica*. Kinetic parameters for the three reversible aromatic aminotransferase reactions catalyzed by this enzyme are presented in Tables 3.4a and 3.4b. In the forward direction, the catabolism of L-tyrosine was clearly the preferred reaction as it produced the highest catalytic efficiency (as indicated by $k_{cat}/K_m$), reaction rate ($V_{max}$) and substrate affinity (indicated by the lowest $K_m$ value). These parameters were significantly different between the three aromatic substrates according to Fisher's protected LSD test ($p \leq 0.05$). In the reverse direction, the highest catalytic efficiency was observed when using 4-hydroxyphenylpyruvate – the conjugate oxo-acid of L-tyrosine.

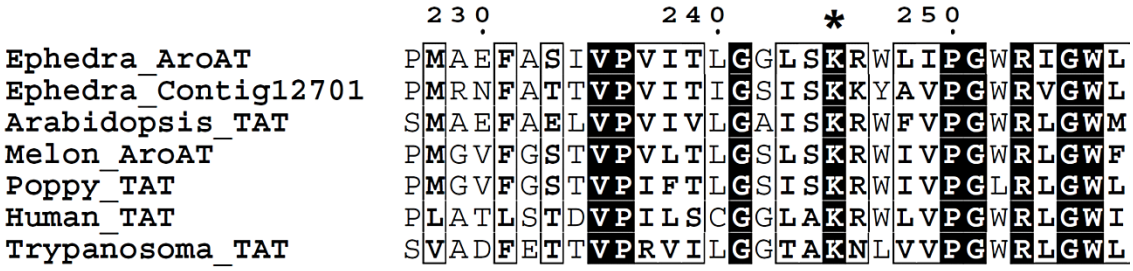


**Figure 4.1 Multiple sequence alignment of aromatic aminotransferase enzymes focusing on the active site lysine residue.** Deduced amino acid sequences for *Ephedra sinica* Contig13244 (AroAT) and Contig12701 were aligned with tyrosine aminotransferases (TAT) from *Arabidopsis thaliana* (GenBank accession number: AED94050.1), *Papaver somniferum* (ADC33123.1), *Homo sapiens* (EAW59231.1) and *Trypanosoma cruzi* (AAA02975.1) as well as an aromatic aminotransferase from musk melon – *Cucumis melo* (ADC45389.1). Clustal Omega (EBI) was used to generate the sequence alignment and BOXSHADE was used to shade the alignment based on similarity. The conserved lysine residue where PLP binds is denoted with an asterisk (*). Note the non-synonymous substitutions from arginine and tryptophan to lysine and tyrosine, respectively after the conserved lysine residue which is only present in Contig12701.

Phenylpyruvate – the conjugate oxo-acid of L-phenylalanine – produced the highest reaction rate. The enzyme had the highest affinity for indole-3-pyruvrate – the conjugate oxo-acid of L-tryptophan. Kinetic parameters were significantly different between the three substrates. It is possible that all three reverse aromatic aminotransferase reactions are catalyzed in the plant by this enzyme as their affinities are all greater than the affinity for L-tyrosine. Catalytic efficiencies for these reactions are also reasonable when compared to previously identified aminotransferases from plants that catalyze reactions involving aromatic substrates (Table 4.1).

*Ephedra sinica* aromatic aminotransferase was more efficient, had greater reaction rates and often had higher affinity when compared to previously identified aromatic aminotransferases from other plant species. Using the data published for *Rosa* aromatic/phenylalanine aminotransferase (Hirata *et al*., 2012), one can calculate the $k_{cat}/K_m$ as 11.47 per mM per second. The catalytic efficiency during the catabolism of L-tyrosine by purified, recombinant 13244 was 32.8 per mM per second, far greater than published values from other plant species (Hirata *et al*., 2012; Lee and Facchini, 2011; Riewe *et al*., 2012). When examining catalytic efficiencies of recombinant 13244 for the anabolism of L-tyrosine, L-phenylalanine and L-tryptophan, it is reasonable to conclude that these reactions may occur *in planta*. The values were much greater than those observed with opium poppy tyrosine aminotransferase involved in benzylisoquinoline biosynthesis (Lee and Facchini, 2011) and were comparable to those seen with *Arabidopsis* tyrosine aminotransferase involved in tocopherol biosynthesis (Riewe *et al*., 2012).

**Table 4.1 Comparison of the kinetic parameters of recently identified plant aromatic aminotransferase enzymes.**

| Enzyme | Substrate | *Forward Reaction* | | | *Reverse Reaction* | | |
|---|---|---|---|---|---|---|---|
| | | $K_m$ (mM) | $V_{max}$ (µmoles/mg/min) | $k_{cat}/K_m$ (/mM/s) | $K_m$ | $V_{max}$ | $k_{cat}/K_m$ |
| *Ephedra* AroAT | Tyrosine | 0.91 ± 0.042 | 42.2 ± 1.07 | 36.1 ± 0.756 | 0.105 ± 0.018 | 0.806 ± 0.05 | 6.06 ± 0.930 |
| | Phenylalanine | 8.67 ± 1.21 | 10.8 ± 1.06 | 0.972 ± 0.043 | 0.364 ± 0.040 | 1.15 ± 0.037 | 2.47 ± 0.202 |
| | Tryptophan | 5.37 ± 1.04 | 4.54 ± 0.224 | 0.672 ± 0.111 | 0.037 ± 0.009 | 0.214 ± 0.038 | 4.51 ± 0.432 |
| Poppy TAT | Tyrosine | 1.82 ± 0.09 | 0.63 ± 0.02 | 0.13 ± 0.07 | Not Reported (NR) | | |
| | Phenylalanine | 6.33 ± 1.09 | 0.19 ± 0.002 | 0.01 ± 0.001 | | | |
| | Tryptophan | 7.83 ± 3.81 | 0.92 ± 0.02 | 0.05 ± 0.002 | | | |
| *Arabidopsis* TAT (Prabhu et al., 2010) | Tyrosine | 0.19 ± 0.16 | 5.0 ± 0.3 | NR | 0.22 ± 0.05 | 23.5 ± 1.0 | NR |
| | Phenylalanine | 0.84 ± 0.2 | NR | | 0.13 ± 0.06 | NR | |
| Rose PheAT | Phenylalanine | 1.47 ± 0.37 | 21.85 ± 0.59 | NR | *Activity 9.7-fold less then Forward | | |
| *Arabidopsis* TAT (Riewe et al., 2012) | Tyrosine | 0.18 ± 0.02 | 1.73 ± 0.05 | 7.72 ± 0.78 | NR | 0.071 ± 0.01 | NR |

Previously, the highest reaction rates were observed during the catabolism of L-phenylalanine by *Rosa* aromatic aminotransferase (Hirata *et al*., 2012) and anabolism of L-tyrosine by *Arabidopsis* tyrosine aminotransferase (Prabhu and Hudson, 2010). However, *Ephedra sinica* aromatic aminotransferase produced an approximately two-fold higher reaction rate of 42.4 µmoles per minute per mg of protein. Substrate affinity observed with *Ephedra sinica* recombinant 13244 was often comparable to the values observed with previously characterized plant aromatic aminotransferases. The $K_m$ values observed during the forward reactions were very similar to those observed with opium poppy tyrosine aminotransferase (Lee and Facchini, 2011) and rose aromatic aminotransferase (Hirata *et al*., 2012). For the reverse reactions, $K_m$ values were very similar to the values observed with *Arabidopsis* tyrosine aminotransferase (Prabhu and Hudson, 2010) which preferred to produce L-tyrosine instead of catabolize it.

Due to their large L-phenylalanine requirements, higher plants may have evolved two pathways for its production. It was believed that L-phenylalanine is produced solely via L-arogenate in higher plants and not via transamination of prephenate due to the fact that prephenate dehydratase (PDT) activity, which produces phenylpyruvate from prephenate, had not been detected in any land plant (Jung *et al*., 1986). Recently, two of the *Arabidopsis* arogenate dehydratase (ADT) enzymes, *AtADT1* and *AtADT2* were shown to have PDT activity *in vitro* (Cho *et al*., 2007) and *in vivo* by complementing yeast mutants that lack PDT activity (Bross *et al*., 2011). A bifunctional PDT and ADT enzyme from rice was recently characterized as well (Yamada *et al*., 2008). Figure 4.2 presents the proposed dual pathways for L-phenylalanine production in higher plants (Yamada *et al*., 2008).

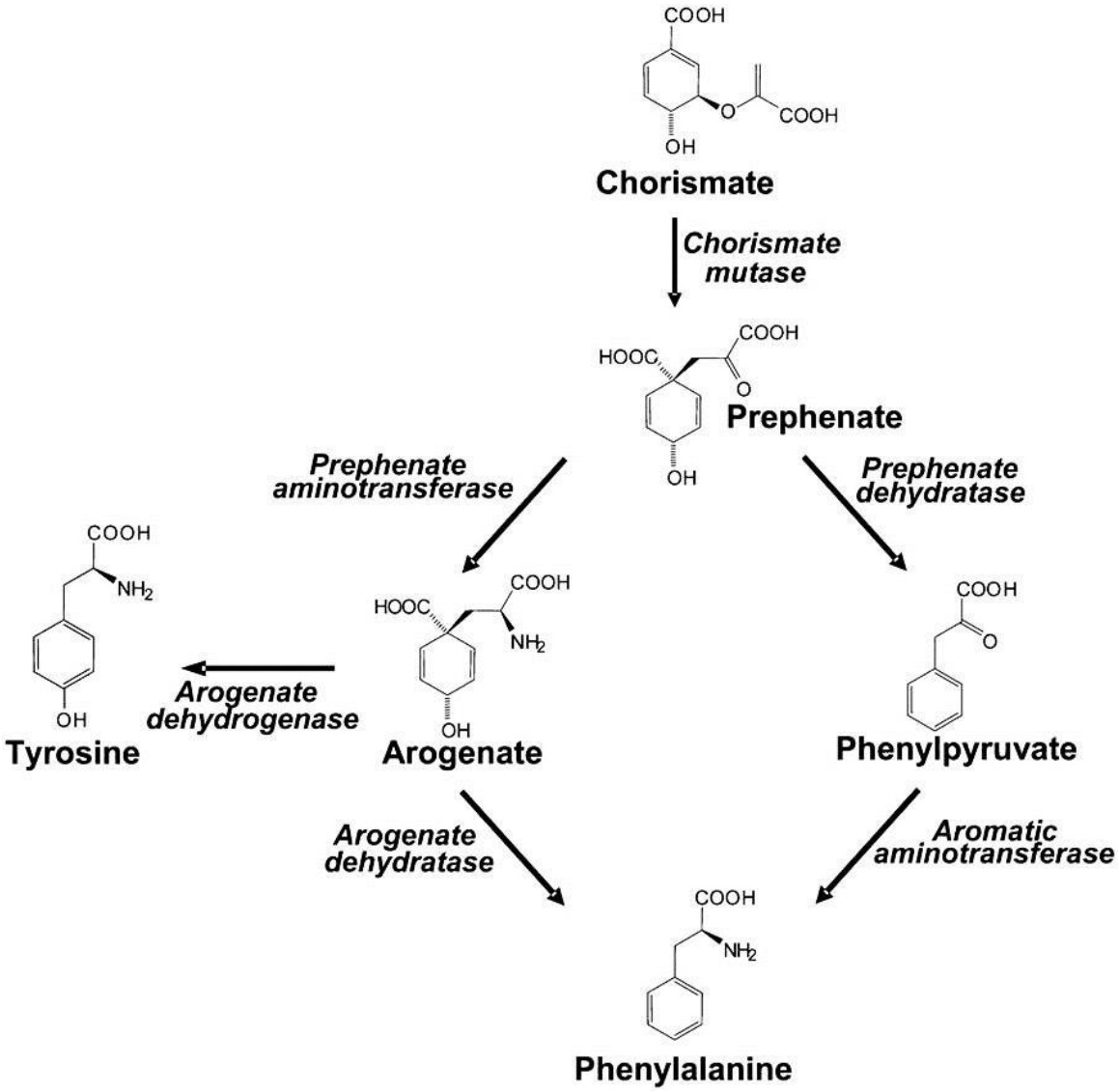


**Figure 4.2 Presumed dual pathways for L-phenylalanine synthesis from chorismate in higher plants.** L-Phenylalanine can be produced via arogenate by arogenate dehydratase or via phenylpyruvate by an aromatic aminotransferase. (Reproduced from Yamada *et al*., 2008 with permission from the American Society of Plant Physiologists)

Transcripts of these bifunctional enzymes exhibiting PDT activity are similar to multiple cDNAs in the *Ephedra sinica* EST database (Table 4.2). Phylogenetic analysis revealed ADT/PDT enzymes group together based on their function. Enzymes which exhibit PDT activity were found in subgroups one and two while enzymes which lack PDT activity were grouped together in subgroup three (Cho *et al.*, 2007). Two cDNAs from the *Ephedra sinica* EST database, Contig16187 and Singlet93640 were grouped together with *AtADT2* and the rice PDT/ADT in subgroup two when *Ephedra* sequences were included in the phylogenetic analysis (Figure 4.3). The other three *Ephedra sinica* PDT candidates were found in subgroup three. This is similar to the ADTs/PDTs from *Arabidopsis*: most do not exhibit PDT activity and are grouped together in subgroup three and two which have PDT activity are found in the other subgroups. Since they are found in subgroup two, it is likely that Contig16187 and Singlet93640 exhibit PDT activity and therefore, the alternative pathway of L-phenylalanine biosynthesis likely exists in *Ephedra*. *Ephedra sinica* aromatic aminotransferase (Contig13244) likely contributes to ephedrine alkaloid biosynthesis by producing, to some extent, the initial precursor of the ephedrine alkaloids, L-phenylalanine.

**Table 4.2 Candidate prephenate dehydratases involved in the conversion of prephenate to phenylpyruvate.**

| tBLASTn Query | | | Top Hits in *Ephedra sinica* EST Collection | | |
|---|---|---|---|---|---|
| ***Accession*** | ***Enzyme*** | ***Organism*** | ***EsUniGene*** | ***Score/Evalue*** | ***Identities/Positives (%)*** |
| AK066428 | Prephenate/Arogenate Dehydratase | Rice | Contig16187 | 453/1.00E-127 | 70/81 |
| | | | Singlet100613 | 452/1.00E-127 | 70/81 |
| | | | Contig11893 | 404/1.00E-112 | 64/79 |
| | | | Singlet93640 | 398/1.00E-111 | 63/76 |
| | | | Contig12662 | 393/1.00E-109 | 62/76 |
| ABD67752.1 | Arogenate/Prephenate Dehydratase1 | *Arabidopsis* | Singlet100613 | 469/1.00E-132 | 71/84 |
| | | | Contig16187 | 468/1.00E-132 | 71/84 |
| | | | Singlet93640 | 419/1.00E-117 | 60/76 |
| | | | Contig11893 | 401/1.00E-112 | 57/73 |
| | | | Singlet25323 | 394/1.00E-109 | 59/74 |
| | | | Contig12662 | 391/1.00E-109 | 55/72 |
| | | | Singlet12017 | 361/1.00E-99 | 57/75 |
| ABD67751.1 | Arogenate/Prephenate Dehydratase2 | *Arabidopsis* | Contig16187 | 466/1.00E-131 | 72/81 |
| | | | Singlet100613 | 466/1.00E-131 | 72/81 |
| | | | Singlet93640 | 422/1.00E-118 | 60/73 |
| | | | Singlet25323 | 400/1.00E-111 | 58/72 |
| | | | Contig11893 | 395/1.00E-110 | 66/79 |
| | | | Contig12662 | 392/1.00E-109 | 63/75 |

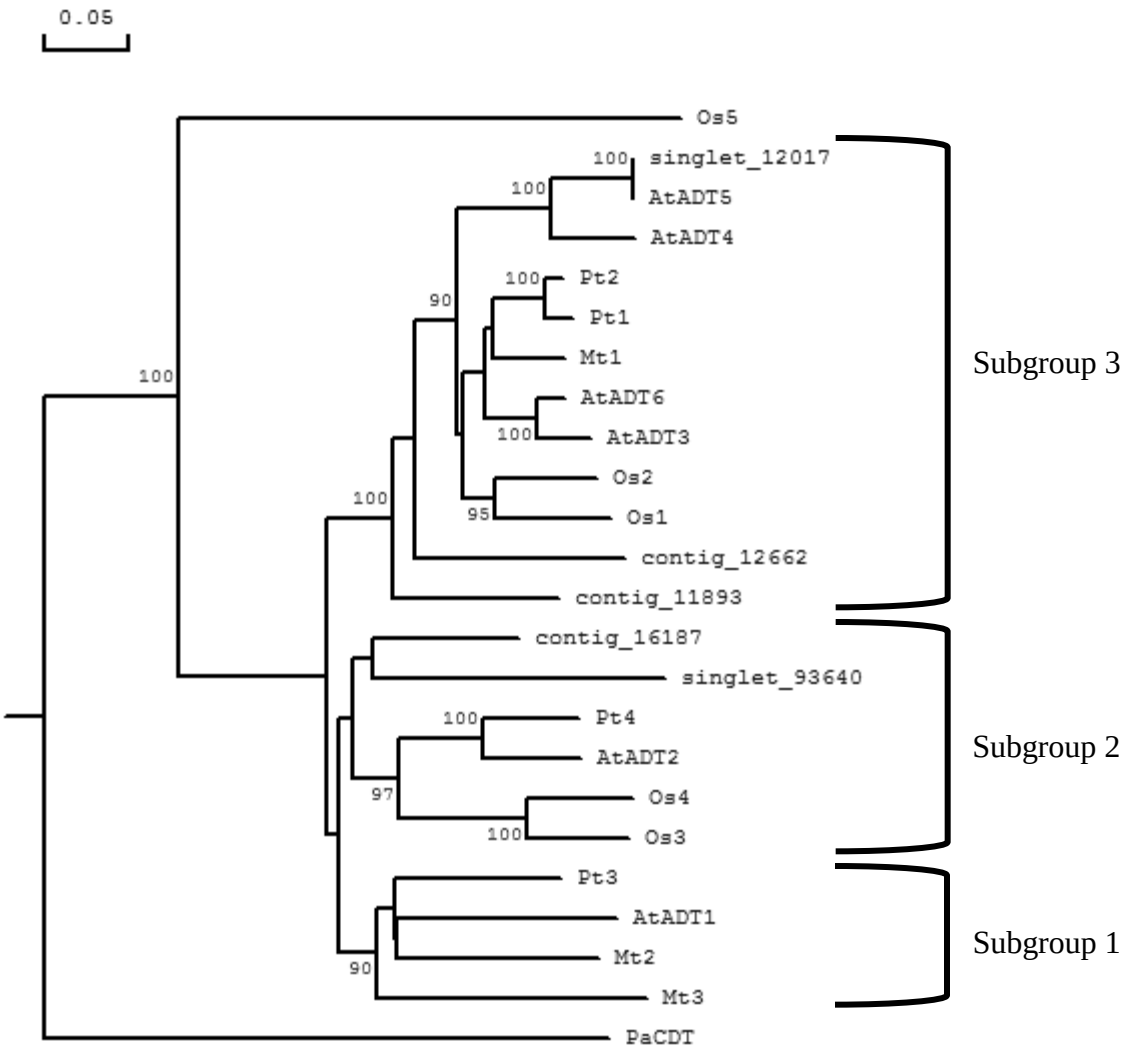


**Figure 4.3 Phylogenetic relationship between selected plant ADT/PDT enzymes.** This rooted phylogenetic tree was generated with DNAMAN using a bootstrap of 1000 and compares amino acid sequences (excluding transit peptide sequences) of various plant ADTs/PDTs. The numbers at the branch points are the bootstrapping values and the horizontal scale indicates sequence divergence. The plant ADT/PDT subgroups are denoted by brackets to the right of the tree. The CDT sequence for *Pseudomonas aeruginosa* (GenBank accession number: AAC08596) was included as an outlier to root the tree. Plant sequences (GenBank accession numbers or EST identification codes in brackets) used were as follows; *Arabidopsis thaliana* sequences: AtADT1 (AY081528), AtADT2 (AY084830), AtADT3 (AY087695), AtADT4 (AY062692), AtADT5 (AY058097), AtADT6 (AY056290); *Medicago truncatula* sequences: Mt1 (TC80102), Mt2 (TC86574), Mt3 (TC87395); *Oryza sativa* sequences: Os1 (AK103609), Os2 (AK066427), Os3 (AK066428), Os4 (AK066644), Os5 (AK099858); *Popular trichocarpa* sequences: Pt1 (EEE86077), Pt2 (EEE75953), Pt3 (EEE97147), Pt4 (EEE87606). Five putative ADTs/PDTs from the *Ephedra sinica* annotated EST database were also included: Contig16187, Contig11893, Singlet93640, Contig 12662 and Singlet12017.

### 4.5 Enzymatic Screening of *Ephedra sinica* Cell-free Extracts

Forward L-tyrosine aminotransferase and cathinone aminotransferase/synthase activities were also screened for in cell-free extracts of young *Ephedra sinica* stem tissue. Forward L-tyrosine aminotransferase activity was detected in both the 30-60% and 60-90% ammonium sulfate fractions confirming the presence of the enzyme in this tissue. Unfortunately, screening for cathinone aminotransferase/synthase activity was unsuccessful by spectrophotometry in cell-free extracts. It may be possible to utilize a GC-MS method to detect cathinone aminotransferase/synthase activity in *Ephedra sinica* cell-free extracts. Benzaldehyde carboxyligase, cathinone reductase and *N*-methyltransferase activities have been detected in cell-free extracts of *Ephedra sinica* stem tissue using GC-MS analysis (Krizevski *et al*., 2012).

### 4.6 Conclusions

In conclusion, the annotated EST database developed in this study formed a functional genomics platform to investigate biosynthesis of the pharmaceutically important ephedrine alkaloids. It marked the first time next-generation sequencing and advanced "-omics" methodologies have been utilized to study ephedrine alkaloid biosynthesis. This deep transcriptome is a crucial resource for gene discovery and should aid in the full characterization of the biochemical processes underlying biosynthesis of these pharmacoactive psychostimulants. This resource also helps to confirm the putative biosynthetic route to the ephedrine alkaloids by identifying a number of gene candidates for each enzymatic reaction in the pathway.

Seventeen candidate cDNAs putatively involved in the conversion of 1-phenylpropane-1,2-dione to (*S*)-cathinone and/or the conversion of phenylpyruvate to L-phenylalanine were successfully cloned and eleven were successfully expressed in a heterologous system. Enzymatic assays based on spectrophotometry were developed for the six aromatic aminotransferase reactions as well as the production of (*S*)-cathinone from 1-phenylpropane-1,2-dione. Recombinant enzymes were assayed for aromatic aminotransferase and cathinone aminotransferase/synthase activity *in vitro*. An *Ephedra sinica* aromatic aminotransferase able to complete all three reversible reactions was discovered and fully characterized. This enzyme is likely involved in ephedrine alkaloid biosynthesis by producing, to some extent, the initial precursor of the pathway, L-phenylalanine. The cathinone aminotransferase/synthase enzyme remained elusive; it was not discovered in successfully expressed candidates from the EST database and its activity could not be detected by spectrophotometry in cell-free extracts of *Ephedra sinica* stem tissue. The remaining candidates that were not expressed may or may not represent or encode the enzyme that produces (*S*)-cathinone. It is possible that the cathinone aminotransferase/synthase transcript has very little sequence similarity to previously characterized aminotransferases and therefore would not be identified through functional genomic analysis. While this study was unsuccessful in discovering the enzyme responsible for the production of (*S*)-cathinone from 1-phenylpropane-1,2-dione, it was still an overall success. The functional genomics platform developed provides an instrumental tool for gene discovery in the pharmaceutically important plant, *Ephedra sinica*. It also serves as a resource for further ephedrine alkaloid research in other plant species to compare, contrast and aid in gene discoveries and pathway elucidation. A highly active aromatic aminotransferase was discovered and fully characterized. It is a

key enzyme involved in aromatic amino acid metabolism and it likely contributes to the production of L-phenylalanine, the initial precursor of the ephedrine alkaloids.

### 4.7 Future Implications

In order to successfully identify the enzyme responsible for the conversion of 1-phenylpropane-1,2-dione to (*S*)-cathinone, the list of gene candidates should be expanded to include all aminotransferase fold types. Also, candidate cDNAs which could not be expressed in *E. coli* should be expressed in an alternative system, possibly by *in vitro* cell-free protein synthesis, so their functions can be characterized. In order to detect cathinone aminotransferase/synthase activity in cell-free *Ephedra sinica* extracts, a GC-MS method of analysis should be developed similar to those used for the detection of benzaldehyde carboxyligase, cathinone reductase and *N*-methyltransferase activity (Krizevski *et al*., 2012).

In order to confirm *Ephedra sinica* aromatic aminotransferase contributes to ephedrine alkaloid biosynthesis, a method of gene knockout or knockdown should be developed. To date, viral induced gene silencing (VIGS) and other gene knockdown methodologies have not been developed in any *Ephedra* species. A VIGS method was successfully utilized to knockdown expression of *Papaver somniferum* tyrosine aminotransferase to show it contributed to benzylisoquinoline alkaloid production (Lee and Facchini, 2011). Biolistic transformation has been successfully used in the conifer species *Pinus radiata* to silence gene expression (Wagner *et al*., 2005; Wagner *et al*., 2007). Successful VIGS has yet to be reported in a conifer species.

Future studies investigating ephedrine alkaloid biosynthesis will benefit greatly from access to the annotated *Ephedra sinica* EST database developed in this study. This resource is instrumental to the discovery of biosynthetic genes involved in the pathway. In the long term, the availability of cloned enzymes of ephedrine alkaloid biosynthesis may allow the development of synthetic biology platforms for the synthesis of amphetamine analogues with known or novel pharmaceutical properties.

# Bibliography

Avenaud, P., Castroviejo, M., Claret, S., Rosenbaum, J., Megraud, F., and Menard, A. (2004). Expression and activity of the cytolethal distending toxin of *Helicobacter hepaticus*. Biochem Biophys Res Commun *318*, 739-745.

Bensky, D., Gamble, A., and Kaptchuk, T., eds. (1986). Chinese Herbal Medicine (Materia Medica. Eastland, Seattle, Washington.)

Boatright, J., Negre, F., Chen, X.L., Kish, C.M., Wood, B., Peel, G., Orlova, I., Gang, D., Rhodes, D., and Dudareva, N. (2004). Understanding *in vivo* benzenoid metabolism in petunia petal tissue. Plant Physiol *135*, 1993-2011.

Bross, C.D., Corea, O.R.A., Kaldis, A., Menassa, R., Bernards, M.A., and Kohalmi, S.E. (2011). Complementation of the *pha2* yeast mutant suggests functional differences for arogenate dehydratases from *Arabidopsis thaliana*. Plant Physiol Biochem *49*, 882-890.

Bruneton, J. (1995). Pharmacognosy, Phytochemistry, Medicinal Plants (Intercept Ltd. Hampshire.)

Chen, S., Luo, H., Li, Y., Sun, Y., Wu, Q., Niu, Y., Song, J., Lv, A., Zhu, Y., Sun, C., *et al.* (2011). 454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in *Panax ginseng*. Plant Cell Reports *30*, 1593-1601.

Cho, M.H., Corea, O.R.A., Yang, H., Bedgar, D.L., Laskar, D.D., Anterola, A.M., Moog-Anterola, F.A., Hood, R.L., Kohalmi, S.E., Bernards, M.A., *et al.* (2007). Phenylalanine biosynthesis in *Arabidopsis thaliana* - Identification and characterization of arogenate dehydratases. J Biol Chem *282*, 30827-30835.

Collier, R.H., and Kohlhaw, G. (1972). Nonidentity of aspartate and aromatic aminotransferase components of transaminase-a in *Escherichia-coli*. Journal of Bacteriology *112*, 365-371.

Cui, J.-F., Zhou, T.-H., Zhang, J.-S., and Lou, Z.-C. (1991). Analysis of alkaloids in Chinese *Ephedra* species by gas chromatographic methods. Phytochemical Analysis *2*, 116-119.

de la Torre, F., Santis, L.D., Suárez, M.F., Crespillo, R., and Cánovas, F.M. (2006). Identification and functional analysis of a prokaryotic-type aspartate aminotransferase: implications for plant amino acid metabolism. The Plant Journal *46*, 414-425.

Desgagne-Penix, I., Khan, M., Schriemer, D., Cram, D., Nowak, J., and Facchini, P. (2010). Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures. BMC Plant Biology *10*, 252-268.

Eliot, A.C., and Kirsch, J.F. (2004). Pyridoxal phosphate enzymes: Mechanistic, structural, and evolutionary considerations. Annual Review of Biochemistry *73*, 383-415.

Facchini, P.J., and De Luca, V. (2008). Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. Plant J *54*, 763-784.

Givan, C.V. (1980). Amino transferases in higher plants. In The biochemistry of plants, B.J. Miflin, ed. (Academic Press), pp. 329-357.

Gonda, I., Bar, E., Portnoy, V., Lev, S., Burger, J., Schaffer, A.A., Tadmor, Y.a., Gepstein, S., Giovannoni, J.J., Katzir, N., *et al.* (2010). Branched-chain and aromatic amino acid catabolism into aroma volatiles in *Cucumis melo* L. fruit. Journal of Experimental Botany *61*, 1111-1123.

Grue-Sørensen, G., and Spenser, I.D. (1989). The biosynthesis of ephedrine. Canadian Journal of Chemistry-Revue Canadienne De Chimie *67*, 998-1009.

Grue-Sørensen, G., and Spenser, I.D. (1993). Biosynthesis of the *Ephedra* alkaloids – Evolution of the C6-C3 skeleton. Journal of the American Chemical Society *115*, 2052-2054.

Grue-Sørensen, G., and Spenser, I.D. (1994). Biosynthetic route to the *Ephedra* alkaloids. Journal of the American Chemical Society *116*, 6195-6200.

Hagel, J.M., Krizevski, R., Kilpatrick, K., Sitrit, Y., Marsolais, F., Lewinsohn, E., and Facchini, P.J. (2011). Expressed sequence tag analysis of khat (*Catha edulis*) provides a putative molecular biochemical basis for the biosynthesis of phenylpropylamino alkaloids. Genetics and Molecular Biology *34*, 640-646.

Hagel, J.M., Krizevski, R., Marsolais, F., Lewinsohn, E., and Facchini, P.J. (2012). Biosynthesis of amphetamine analogs in plants. Trends in Plant Science *17*, 404-412.

Haller, C.A., and Benowitz, N.L. (2000). Adverse cardiovascular and central nervous system events associated with dietary supplements containing *Ephedra* alkaloids. New England Journal of Medicine *343*, 1833-1838.

Hirata, H., Ohnishi, T., Ishida, H., Tomida, K., Sakai, M., Hara, M., and Watanabe, N. (2012). Functional characterization of aromatic amino acid aminotransferase involved in 2-phenylethanol biosynthesis in isolated rose petal protoplasts. J Plant Physiol *169*, 444-451.

Hirotsu, K., Goto, M., Okamoto, A., and Miyahara, I. (2005). Dual substrate recognition of aminotransferases. Chemical Record *5*, 160-172.

Invitrogen (2010). Expressway milligram cell-free *E. coli* expression system instruction manual (Invitrogen Life Technologies).

Jung, E., Zamir, L.O., and Jensen, R.A. (1986). Chloroplasts of higher-plants synthesize L-phenylalanine via L-arogenate. Proc Natl Acad Sci U S A *83*, 7231-7235.

Khatoon, S., Srivastava, M., Rawat, A.K.S., and Mehrotra, S. (2005). HPTLC method for chemical standardization of *Sida* species and estimation of the alkaloid ephedrine. Jpc-Journal of Planar Chromatography-Modern Tlc *18*, 364-367.

Kingston, R.E. (2001). Preparation and Analysis of RNA. In Current Protocols in Molecular Biology (John Wiley & Sons, Inc.).

Kirsch, J.F., Eichele, G., Ford, G.C., Vincent, M.G., Jansonius, J.N., Gehring, H., and Christen, P. (1984). Mechanism of action of aspartate-aminotransferase proposed on the basis of its spatial structure. J Mol Biol *174*, 497-525.

Klammt, C., Schwarz, D., Löhr, F., Schneider, B., Dötsch, V., and Bernhard, F. (2006). Cell-free expression as an emerging technique for the large scale production of integral membrane protein. FEBS Journal *273*, 4141-4153.

Kochhar, S., and Christen, P. (1992). Mechanism of racemization of amino-acids by aspartate-aminotransferase. Eur J Biochem *203*, 563-569.

Kochhar, S., Finlayson, W.L., Kirsch, J.F., and Christen, P. (1987). The stereospecific labilization of the C-4' PRO-S hydrogen of pyridoxamine 5'-phosphate is abolished in (lys258- ala) aspartate-aminotransferase. J Biol Chem *262*, 11446-11448.

Krizevski, R., Bar, E., Shalit, O., Levy, A., Hagel, J.M., Kilpatrick, K., Marsolais, F., Facchini, P.J., Ben-Shabat, S., Sitrit, Y., *et al.* (2012). Benzaldehyde is a precursor of phenylpropylamino alkaloids as revealed by targeted metabolic profiling and comparative biochemical analyses in *Ephedra* spp. Phytochemistry *81*, 71-79.

Krizevski, R., Bar, E., Shalit, O., Sitrit, Y., Ben-Shabat, S., and Lewinsohn, E. (2010). Composition and stereochemistry of ephedrine alkaloids accumulation in *Ephedra sinica* Stapf. Phytochemistry *71*, 895-903.

Krizevski, R., Dudai, N., Bar, E., Dessow, I., Ravid, U., and Lewinsohn, E. (2008). Quantitative stereoisomeric determination of phenylpropylamino alkaloids in khat (*Catha edulis* Forsk.) using chiral GC-MS. Israel Journal of Plant Sciences *56*, 207-213.

Krizevski, R., Dudai, N., Bar, E., and Lewinsohn, E. (2007). Developmental patterns of phenylpropylamino alkaloids accumulation in khat (*Catha edulis*, Forsk.). Journal of Ethnopharmacology *114*, 432-438.

Kuramitsu, S., Inoue, Y., Tanase, S., Morino, Y., and Kagamiyama, H. (1987). Substitution of an arginyl residue for the active-site lysyl residue (lys258) of aspartate-aminotransferase. Biochem Biophys Res Commun *146*, 416-421.

Lee, E.-J., and Facchini, P.J. (2011). Tyrosine Aminotransferase Contributes to Benzylisoquinoline Alkaloid Biosynthesis in Opium Poppy. Plant Physiology *157*, 1067-1078.

Leete, E. (1958). Biogenesis of D-norpseudo ephedrine in *Catha-edulis*. Chemistry & Industry, 1088-1089.

Leonhart, M.M. (2008). Proposed Revised Assessment of Annual Needs for the List I Chemicals Ephedrine, Pseudoephedrine, and Phenylpropanolamine for 2008. http://www.deadiversion.usdoj.gov/fed_regs/quotas/2008/fr0623.htm.

Lewis, W.H., and Elvin-Lewis, M.P.F. (1977). Medical botany, plants affecting man's health (Wiley Interscience, USA).

Liepman, A.H., and Olsen, L.I. (2004). Genomic analysis of aminotransferases in *Arabidopsis thaliana*. Crit Rev Plant Sci *23*, 73-89.

Maeda, H., and Dudareva, N. (2012). The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. Annual Review of Plant Biology, Vol 63 *63*, 73-105.

Maeda, H., Yoo, H., and Dudareva, N. (2011). Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. Nature Chemical Biology *7*, 19-21.

Mehta, P.K., Hale, T.I., and Christen, P. (1993). Aminotransferases - Demonstration of homology and division into evolutionary subgroups. Eur J Biochem *214*, 549-561.

Meyer, E., Aglyamova, G.V., Wang, S., Buchanan-Carter, J., Abrego, D., Colbourne, J.K., Willis, B.L., and Matz, M.V. (2009). Sequencing and *de novo* analysis of a coral larval transcriptome using 454 GS-FLX. BMC Genomics *10*, 219-236.

Montemartini, M., Santome, J.A., Cazzulo, J.J., and Nowicki, C. (1993). Purification and partial structural and kinetic characterization of tyrosine aminotransferase from epimastigotes of *Trypanosoma-cruzi*. Biochem J *292*, 901-906.

NEB (2012). PURExpress *In Vitro* Protein Synthesis Instruction Manual (New England BioLabs Inc.).

Nierop Groot, M.N., and De Bont, J.A.M. (1998). Conversion of phenylalanine to benzaldehyde initiated by an aminotransferase in *Lactobacillus plantarum*. Applied and Environmental Microbiology *64*, 3009-3013.

Okada, T., Mikage, M., and Sekita, S. (2008). Molecular Characterization of the Phenylalanine Ammonia-Lyase from *Ephedra sinica*. Biological & Pharmaceutical Bulletin *31*, 2194-2199.

Oshio, H., Tsukui, M., and Matsuoka, T. (1978). Isolation of L-ephedrine from *Pinelliae* tuber. Chemical & Pharmaceutical Bulletin *26*, 2096-2097.

Prabhu, P.R., and Hudson, A.O. (2010). Identification and Partial Characterization of an L-Tyrosine Aminotransferase (TAT) from *Arabidopsis thaliana*. Biochemistry research international *2010*, 549572.

Qiagen (2008). EasyXpress Protein Synthesis Handbook.

Riewe, D., Koohi, M., Lisec, J., Pfeiffer, M., Lippmann, R., Schmeichel, J., Willmitzer, L., and Altmann, T. (2012). A tyrosine aminotransferase involved in tocopherol synthesis in *Arabidopsis*. Plant J *71*, 850-859.

Rydin, C., and Korall, P. (2009). Evolutionary relationships in *Ephedra* (Gnetales), with implications for seed plant phylogeny. International Journal of Plant Sciences *170*, 1031-1043.

Shibata, S., Imaseki, I., and Yamazaki, M. (1957). Phytochemical investigation on cultivation of medicinal plants. XII. On the alkaloid biogenesis in *Ephedra.* (3). Pharmaceutical bulletin *5*, 71-73.

Sigma (2012). *In Vitro* Director PCR System (Sigma-Aldrich).

Sivaraman, S., and Kirsch, J.F. (2006). The narrow substrate specificity of human tyrosine aminotransferase – the enzyme deficient in tyrosinemia type II. FEBS Journal *273*, 1920-1929.

Stepanova, A.N., Robertson-Hoyt, J., Yun, J., Benavente, L.M., Xie, D.-Y., Doležal, K., Schlereth, A., Jürgens, G., and Alonso, J.M. (2008). TAA1-Mediated Auxin Biosynthesis Is Essential for Hormone Crosstalk and Plant Development. Cell *133*, 177-191.

Tabor, S. (2001). Expression Using the T7 RNA Polymerase/Promoter System. In Current Protocols in Molecular Biology (John Wiley & Sons, Inc.).

ThermoScientific (2011). 1-Step Human High-Yield Mini IVT Kit Instructions (Pierce Biotechnology, Rockford, lL, USA).

Udvardi, M.K., and Kahn, M.L. (1991). Isolation and analysis of a cDNA clone that encodes an alfalfa (*Medicago sativa*) aspartate aminotransferase. Molecular and General Genetics MGG *231*, 97-105.

Wagner, A., Phillips, L., Narayan, R.D., Moody, J.M., and Geddes, B. (2005). Gene silencing studies in the gymnosperm species *Pinus radiata*. Plant Cell Reports *24*, 95-102.

Wagner, A., Ralph, J., Akiyama, T., Flint, H., Phillips, L., Torr, K., Nanayakkara, B., and Kiri, L.T. (2007). Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in *Pinus radiata*. Proc Natl Acad Sci U S A *104*, 11856-11861.

Wang, C.-S., and Vodkin, L. (1994). Extraction of RNA from tissues containing high levels of procyanidins that bind RNA. Plant Molecular Biology Reporter *12*, 132-145.

Yamada, T., Matsuda, F., Kasai, K., Fukuoka, S., Kitamura, K., Tozawa, Y., Miyagawa, H., and Wakasa, K. (2008). Mutation of a rice gene encoding a phenylalanine biosynthetic enzyme results in accumulation of phenylalanine and tryptophan. Plant Cell *20*, 1316-1329.

Yamasaki, K., Sankawa, U., and Shibata, S. (1969). Biosynthesis of ephedrine in *Ephedra-distachya-g* participation of 6 carbon 1 carbon unit. Tetrahedron Letters *47*, 4099-4102.

Yamasaki, K., Tamaki, T., Uzawa, S., Sankawa, U., and Shibata, S. (1973). Biosynthesis of natural-products .8. Participation of C6-C1 unit in biosynthesis of ephedrine in *Ephedra*. Phytochemistry *12*, 2877-2882.

Zerbino, D.R., and Birney, E. (2008). Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res *18*, 821-829.

Ziegler, J., and Facchini, P.J. (2008). Alkaloid biosynthesis: Metabolism and trafficking. In Annual Review of Plant Biology (Palo Alto, Annual Reviews), pp. 735-769.

Ziegler, J., Facchini, P.J., Geissler, R., Schmidt, J., Ammer, C., Kramell, R., Voigtlaender, S., Gesell, A., Pienkny, S., and Brandt, W. (2009). Evolution of morphine biosynthesis in opium poppy. Phytochemistry *70*, 1696-1707.

# Appendices

**Appendix A1 PCR primers used to clone *Ephedra sinica* aromatic aminotransferase candidates.**

| *EsUniGene* | *Primer Name* | *Primer Sequence* |
|---|---|---|
| Contig4103 | Contig4103fwd | CTACGGATCCATGGCAAATCTTTTGTCGAATTCTCAACC |
| | Contig4103rev | CTACCTGCAGTTAGACCAAGGTCTTTTCCGTACATGG |
| Contig12559 | Contig12559f | GGTGATGATGATGACAAGATGGAGATGGAAGCTCTGAATC |
| | Contig12559r | GGAGATGGGAAGTCATTAATCCCCTGCCACACCA |
| Contig13244 | Contig13244fwd | CTACGGATCCATGGCAAGCAATGGGGAATGG |
| | Contig13244rev | CTACGGTACCTTACAAATAAGATGACCTGCAACAGAAAGATTC |
| Singlet18529 | Singlet18529fwd | CTACGGTACCATGGATCATCTCTCAAACCATGCTGC |
| | Singlet18529rev | CTACAAGCTTTTACAAACTCCTAACATGTCTGCTACAAAATGC |
| Contig12701 | Contig12701fwd | CTACGGATCCATGGTGAAGTGGAATTTTAAGCCAAGC |
| | Contig12701rev | CTACAAGCTTCTATAAACTCCTATAATGTCTGCGACAGAATGC |
| Contig19156 | 19156fwdRE | CTACGGATCCATGGGGTCTTCTAGGTTTTTAGTG |
| | 19156revRE | CTACGTCGACTTACTCTATAGTTGAAAGCCTGTCAA |
| Contig901 | 901fwdRE | CTACGGATCCATGATGTACCCATCAACAATAATAATGC |
| | 901revRE | CTACGTCGACTTAGTGGAGCACAAGTTCTTCAAG |
| Contig4535 | 4535fwdRE | CTACGGATCCATGGATAACATGAGCCTCTTAGG |
| | 4535revRE | CTACGTCGACTCAAGCACCAAACACGAGCT |
| Contig18610 | 18610fwdRE | CTACGGATCCATGCTCCTTCAAAGGCCGC |
| | 18610revRE | CTACGTCGACCTAGGCATGACAAGCCAG |
| Contig287 | Contig287f | GGTGATGATGATGACAAGATGGGCAGCAAACCGTTG |
| | Contig287r | GGAGATGGGAAGTCACCTATTGTCAGCATATTTGGC |
| Contig16680 | 16680fwdRE | CTACGGATCCATGTGCAAGGTTTATGTGATCTTC |
| | 16680revRE | CTACGTCGACTCAACTTTCTCTCAAGACCAAAG |
| Contig29549 | 29549fwdRE | CTACGGATCCATGCGGAGATTTGTCGCAG |
| | 29549revRE | CTACGTCGACTCAAAACCTAAATGTACCTTTAGATACAG |
| Singlet4072 | Singlet4072f | GGTGATGATGATGACAAGATGGCCAAAACTCTGAAATCATATTG |
| | Singlet4072r | GGAGATGGGAAGTCATGGACTGATTCCTTTATTTACC |
| Contig19895 | Contig19895f1 | GGTGATGATGATGACAAGATGGCGTGCATATCCTGG |
| | Contig19895r | GGAGATGGGAAGTCATTAAGCGGCTTTCCTGTAAAG |
| Contig11412 | 11412fwdRE | CTACGGATCCATGAATGACATGAATGATTCGAGTTTAG |
| | 11412revRE | CTACGTCGACCTATATCTTGGTCACTGCAGC |
| Contig22464 | Contig22464f1 | GGTGATGATGATGACAAGATGGCGACGTTTACTGCTTC |
| | Contig22464r | GGAGATGGGAAGTCATTAATTGACATTGTGGAAAGAATCAATG |
| Contig798 | Contig798f | GGTGATGATGATGACAAGATGGAGTATCTTCCTATGGGAG |
| | Contig798r | GGAGATGGGAAGTCACTATGCCTTTGTAACTTCATGGAT |
| Contig12253 | Contig29549f | GGTGATGATGATGACAAGATGCGGAGATTTGTCGCAG |
| | Contig12253r | GGAGATGGGAAGTCAATCACGGAATTCATCCATAAAAC |

**Appendix A2 Comparison of *Ephedra sinica* aromatic aminotransferase candidate clone sizes. Predicted mass includes the corresponding fusion or affinity tag.**

| *EsUniGene* | *Number of Base Pairs* | *Number of Amino Acids* | *Predicted Mass (kDa)* |
|---|---|---|---|
| Contig4103 | 1422 | 473 | 52.4 |
| Contig12559 | 1245 | 414 | 48.4 |
| Contig13244 | 1236 | 411 | 46.7 |
| Singlet18529 | 1293 | 429 | 49.3 |
| Contig12701 | 1245 | 414 | 47.5 |
| Contig19156 | 1581 | 526 | 86.4 |
| Contig901 | 1587 | 528 | 86.5 |
| Contig4535 | 1473 | 490 | 82.5 |
| Contig18610 | 1272 | 423 | 73.8 |
| Contig287 | 1428 | 475 | 56.1 |
| Contig16680 | 1023 | 340 | 64.5 |
| Contig29549 | 1044 | 347 | 65.1 |
| Singlet4072 | 1203 | 400 | 46.9 |
| Contig19895 | 1395 | 464 | 53.3 |
| Contig11412 | 1587 | 528 | 85.3 |
| Contig22464 | 1395 | 464 | 54.0 |
| Contig798 | 1014 | 337 | 40.9 |
| Contig12253 | 1578 | 526 | 60.9 |

**Appendix A3 Permission to use Figure 1.1 from Elsevier**

This is a License Agreement between Korey Kilpatrick ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

| | |
|---|---|
| Supplier | Elsevier Limited<br>The Boulevard, Langford Lane<br>Kidlington, Oxford, OX5 1GB, UK |
| Registered Company Number | 1982084 |
| Customer name | Korey Kilpatrick |
| Customer address | 1391 Sandford Street<br>London, ON N5V4T3 |
| License number | 3018841407320 |
| License date | Oct 30, 2012 |
| Licensed content publisher | Elsevier |
| Licensed content publication | Trends in Plant Science |
| Licensed content title | Biosynthesis of amphetamine analogs in plants |
| Licensed content author | Jillian M. Hagel, Raz Krizevski, Frédéric Marsolais, Efraim Lewinsohn, Peter J. Facchini |
| Licensed content date | July 2012 |
| Licensed content volume number | 17 |
| Licensed content issue number | 7 |
| Number of pages | 9 |
| Start Page | 404 |
| End Page | 412 |
| Type of Use | reuse in a thesis/dissertation |
| Portion | figures/tables/illustrations |

| | |
|---|---|
| Number of figures/tables/ illustrations | 1 |
| Format | electronic |
| Are you the author of this Elsevier article? | No |
| Will you be translating? | No |
| Order reference number | N/A |
| Title of your thesis/dissertation | Functional genomic investigation of aromatic aminotransferases involved in ephedrine alkaloid biosynthesis in *Ephedra sinica* (Stapf) |
| Expected completion date | Dec 2012 |
| Estimated size (number of pages) | 75 |
| Elsevier VAT number | GB 494 6272 12 |
| Permissions price | 0.00 USD |
| VAT/Local Sales Tax | 0.0 USD / 0.0 GBP |
| Total | 0.00 USD |

## Appendix A4 Permission to use Figure 4.1 from American Society of Plant Physiologists

**Order detail ID:** 63247961
**Order License Id:** 3052551016904
**ISSN:** 1532-298X
**Publication Type:** e-Journal
**Volume:** 20
**Issue:** 5
**Start page:** 1316
**Publisher:** AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS
**Author/Editor:** American Society of Plant Physiologists

**Permission Status: Granted**
**Permission type:** Republish or display content
**Type of use:** Republish in a thesis/dissertation

| | |
|---|---|
| Requestor type | Academic institution |
| Format | Electronic |
| Portion | chart/graph/table/figure |
| Number of charts/graphs/tables/figures | 1 |
| Title or numeric reference of the portion(s) | Mutation of a Rice Gene Encoding a Phenylalanine Biosynthetic Enzyme Results in Accumulation of Phenylalanine and Tryptophan - Figure 1. |
| Editor of portion(s) | N/A |
| Author of portion(s) | Tetsuya Yamada, Fumio Matsuda, Koji Kasai, Shuichi Fukuoka, Keisuke Kitamura, Yuzuru Tozawa, Hisashi Miyagawa, and Kyo Wakasaa |
| Volume of serial or monograph | 20 |
| Page range of portion | 1316-1329 |
| Publication date of portion | May 2008 |
| Rights for | Main product |
| Duration of use | 0 - 5 years |
| Creation of copies for the disabled | no |
| With minor editing privileges | no |
| For distribution to | Worldwide |
| In the following language(s) | Original language of publication |

| | |
|---|---|
| With incidental promotional use | no |
| Lifetime unit quantity of new product | 0 to 499 |
| Made available in the following markets | Internet |
| The requesting person/organization | Korey Kilpatrick |
| Order reference number | |
| Author/Editor | Korey Kilpatrick |
| The standard identifier | N/A |
| The proposed price | $0 |
| Title | FUNCTIONAL GENOMIC INVESTIGATION OF AROMATIC AMINOTRANSFERASES INVOLVED IN EPHEDRINE ALKALOID BIOSYNTHESIS IN EPHEDRA SINICA (STAPF) |
| Publisher | UWO Thesis Database |
| Expected publication date | Dec 2012 |
| Estimated size (pages) | 85 |

# *Curriculum Vitae*

**Name:** Korey Kilpatrick

**Post-secondary Education and Degrees:**

2010 – 2012 M.Sc. Biology – Cell and Molecular Biology
The University of Western Ontario
London, Ontario, Canada

2006 – 2010 B.Sc. Biology – Honors Specialization in Genetics
The University of Western Ontario
London, Ontario, Canada

**Honours and Awards:**

2010 – 2012 Western Graduate Research Scholarship

2009 – 2010 Dean's Honor List – The University of Western Ontario

**Related Work Experience:**

2010 – 2013 Teaching Assistant – First Year Biology Laboratories and Tutorials
The University of Western Ontario
London, Ontario, Canada

2010 – 2013 Research Affiliate Program (RAP Student)
Agriculture and Agri-Food Canada – Southern Crop Protection and Food Research Centre
London, Ontario, Canada

**Publications:**

Raz Krizevski, Einat Bar, Or Shalit, Assaf Levi, Jillian M. Hagel, **Korey Kilpatrick**, Frédéric Marsolais, Shimon Ben-Shabat, Yaron Sitrit and Efraim Lewinsohn. (2012). Benzaldehyde is a precursor of phenylpropylamino alkaloids as revealed by targeted metabolic profiling and comparative biochemical analyses of *Ephedra* spp. Phytochemistry *81*: 71 – 79.

Jillian M. Hagel, Raz Krizevski, **Korey Kilpatrick**, Yaron Sitrit, Frédéric Marsolais, Efraim Lewinsohn and Peter J. Facchini. (2011). Expressed sequence tag analysis of khat (*Catha edulis*) provides a putative molecular biochemical basis for the biosynthesis of phenylpropylamino alkaloids. Genet. Mol. Biol. *34*: 640 – 646.